MapReduce Service

User Guide

Issue 01

Date 2026-07-30



HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2026. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Cloud Computing Technologies Co., Ltd.

Trademarks and Permissions


and other Huawei trademarks are the property of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 Product Introduction....1
1.1 What Is MRS?....1
1.2 Advantages....4
1.3 Application Scenarios....8
1.4 MRS Cluster Version Overview....11
1.5 List of MRS Component Versions....11
1.6 Components....13
1.6.1 Alluxio....14
1.6.2 ClickHouse....14
1.6.3 CDL....19
1.6.4 DBService....21
1.6.5 Doris....23
1.6.5.1 Basic Principles....23
1.6.5.2 Relationship with Other Components....26
1.6.6 Flink....26
1.6.6.1 Flink Basic Principles....27
1.6.6.2 Flink HA Solution....32
1.6.6.3 Relationships Between Flink and Other Components....34
1.6.6.4 Flink Sliding Window Enhancement....35
1.6.6.5 Flink Job Pipeline Enhancement....37
1.6.6.6 Flink Stream SQL Join Enhancement....42
1.6.6.7 Flink CEP in SQL Enhancement....43
1.6.7 Flume....45
1.6.7.1 Flume Basic Principles....45
1.6.7.2 Relationships Between Flume and Other Components....48
1.6.7.3 Flume Enhanced Open Source Features....49
1.6.8 Guardian....49
1.6.9 HBase....50
1.6.9.1 HBase Basic Principles....50
1.6.9.2 HBase HA Solution....56
1.6.9.3 Relationship Between HBase and Other Components....57
1.6.9.4 HBase Enhanced Open Source Features....58
1.6.10 HDFS....65

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.10.1 HDFS Basic Principles 65
1.6.10.2 HDFS HA Solution 68
1.6.10.3 Relationship Between HDFS and Other Components 70
1.6.10.4 HDFS Enhanced Open Source Features 72
1.6.11 HetuEngine 79
1.6.11.1 HetuEngine Basic Principles 79
1.6.11.2 Relationships Between HetuEngine and Other Components 82
1.6.12 Hive 82
1.6.12.1 Hive Basic Principles 83
1.6.12.2 Hive CBO Principles 86
1.6.12.3 Relationship Between Hive and Other Components 90
1.6.12.4 Enhanced Open Source Feature 90
1.6.13 Hudi 92
1.6.14 Hue 94
1.6.14.1 Hue Basic Principles 94
1.6.14.2 Relationship Between Hue and Other Components 96
1.6.14.3 Hue Enhanced Open Source Features 98
1.6.15 Impala 98
1.6.16 IoTDB 100
1.6.16.1 IoTDB Basic Principles 100
1.6.16.2 IoTDB Enhanced Open Source Features 103
1.6.17 JobGateway 103
1.6.18 Kafka 104
1.6.18.1 Kafka Basic Principles 104
1.6.18.2 Relationship Between Kafka and Other Components 107
1.6.18.3 Kafka Enhanced Open Source Features 108
1.6.19 KrbServer and LdapServer 108
1.6.19.1 KrbServer and LdapServer Principles 108
1.6.19.2 KrbServer and LdapServer Enhanced Open Source Features 112
1.6.20 Kudu 113
1.6.21 Loader 114
1.6.21.1 Loader Basic Principles 114
1.6.21.2 Relationship Between Loader and Other Components 116
1.6.21.3 Loader Enhanced Open Source Features 116
1.6.22 Manager 117
1.6.22.1 Manager Basic Principles 117
1.6.22.2 Manager Key Features 120
1.6.23 MapReduce 122
1.6.23.1 MapReduce Basic Principles 122
1.6.23.2 Relationship Between MapReduce and Other Components 123
1.6.23.3 MapReduce Enhanced Open Source Features 123
1.6.24 MemArtsCC 126

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.24.1 MemArtsCC Basic Principles 126
1.6.24.2 Relationships Between MemArtsCC and Other Components 128
1.6.25 Oozie 129
1.6.26 OpenTSDB 130
1.6.27 Presto 131
1.6.28 Ranger 132
1.6.28.1 Ranger Basic Principles 132
1.6.28.2 Relationship Between Ranger and Other Components 134
1.6.29 Spark 135
1.6.29.1 Spark Basic Principles 135
1.6.29.2 Spark HA Solution 151
1.6.29.3 Relationship Between Spark and Other Components 157
1.6.29.4 Spark Enhanced Open Source Features 161
1.6.30 Spark2x 164
1.6.30.1 Spark2x Basic Principles 164
1.6.30.2 Spark2x Multi-active Instance 179
1.6.30.3 Spark2x Multi-tenant 182
1.6.30.4 Relationship Between Spark2x and Other Components 185
1.6.30.5 Spark2x Open Source New Features 190
1.6.30.6 Optimizing Spark SQL Query of Data of Multiple Sources 190
1.6.31 Storm 193
1.6.31.1 Storm Basic Principles 193
1.6.31.2 Relationship Between Storm and Other Components 197
1.6.31.3 Storm Enhanced Open Source Features 198
1.6.32 Tez 199
1.6.33 YARN 200
1.6.33.1 YARN Basic Principles 200
1.6.33.2 YARN HA Solution 205
1.6.33.3 Relationship Between YARN and Other Components 206
1.6.33.4 Yarn Enhanced Open Source Features 209
1.6.34 ZooKeeper 217
1.6.34.1 ZooKeeper Basic Principles 217
1.6.34.2 Relationships Between ZooKeeper and Other Components 220
1.6.34.3 ZooKeeper Enhanced Open Source Features 224
1.7 Functions 227
1.7.1 Cluster Lifecycle Management 227
1.7.2 Enterprise Project Management 229
1.7.3 Job Management 229
1.7.4 Metadata Management 230
1.7.5 Managing Multi-Tenancy Resources 230
1.7.6 Easy Access to Web UIs of Components 232
1.7.7 Cluster Management 232

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.7.7.1 Cluster Scaling 232
1.7.7.2 Creating Task Nodes 233
1.7.7.3 Auto Scaling 233
1.7.7.4 Isolating Nodes 235
1.7.7.5 Scaling Up Master Node Specifications 235
1.7.7.6 Managing Node Labels 235
1.7.7.7 Node Bootstrap Actions 236
1.7.7.8 Cluster Status Notification 236
1.7.8 Cluster O&M 237
1.7.9 MRS Security Hardening 239
1.7.10 MRS Reliability Enhancement 241
1.8 Constraints 242
1.9 Permission Management 262
1.10 Related Services 269
1.11 Quota Description 270
1.12 Common Concepts 271
2 Preparations 275
2.1 Creating an IAM User and Granting MRS Permissions 275
2.2 Creating a Custom Policy for MRS 280
3 Quick Start 286
3.1 Creating and Using a Hadoop Cluster for Offline Analysis 286
3.2 Creating and Using a Kafka Cluster for Stream Processing 291
3.3 Creating and Using an HBase Cluster for Offline Query 297
3.4 Creating and Using a ClickHouse Cluster for Columnar Store 302
3.5 Creating and Using an MRS Cluster Requiring Security Authentication 306
4 MRS Cluster Planning 314
4.1 Service Selection 314
4.1.1 MRS Service Scenarios 314
4.1.2 MRS Cluster Node Types 323
4.1.3 MRS Cluster Node Specifications 324
4.2 MRS Cluster Deployment 325
4.2.1 Overview 325
4.2.2 Kerberos Authentication for MRS Clusters 335
4.2.3 Disk Partitions of an MRS Cluster Node 336
5 Buying MRS Clusters 340
5.1 Quickly Creating an MRS Cluster 340
5.2 Creating a Custom MRS Cluster 346
6 Installing an MRS Cluster Client 362
6.1 Installing a Client (MRS 3.*x*) 362
6.2 Installing a Client (MRS 2.x or Earlier) 378

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

7 Submitting an MRS Job 384
7.1 MRS Job Introduction 384
7.2 Uploading Application Data to an MRS Cluster 387
7.3 Submitting a Job in an MRS Cluster 390
7.3.1 Running a MapReduce Job 390
7.3.2 Running a SparkSubmit Job 395
7.3.3 Running a Hive SQL Job 401
7.3.4 Running a Spark SQL Job 406
7.3.5 Running a Flink Job 412
7.4 Viewing MRS Job Details and Logs 420
8 Managing Clusters 424
8.1 MRS Cluster Overview 424
8.2 Accessing MRS Manager 426
8.3 Introduction to MRS Manager 433
8.4 Managing an MRS Cluster 436
8.4.1 Viewing Basic Information About an MRS Cluster 436
8.4.2 Checking the Running Status of an MRS Cluster 438
8.4.3 Starting and Stopping an MRS Cluster 441
8.4.4 Restarting an MRS Cluster 442
8.4.5 Exporting MRS Cluster Configuration Parameters 453
8.4.6 Synchronizing the MRS Cluster Configuration 454
8.4.7 Deleting an MRS Cluster 455
8.4.8 Changing the VPC Subnet of an MRS Cluster 455
8.4.9 Managing the Security Group of an MRS Cluster 458
8.4.10 Modifying the OMS Service Configuration 460
8.4.11 Modifying MRS Manager Routing Table 462
8.5 Managing MRS Cluster Components 465
8.5.1 Checking the Running Status of an MRS Cluster Component 465
8.5.2 Starting and Stopping an MRS Cluster Component 471
8.5.3 Restarting an MRS Cluster Component 473
8.5.4 Adding and Deleting an MRS Cluster Component 487
8.5.5 Modifying the Configuration Parameters of an MRS Cluster Component 490
8.5.6 Viewing the Modified Component Configuration Parameters of an MRS Cluster 493
8.5.7 Synchronizing MRS Component Configuration Parameters 495
8.5.8 Adding Custom MRS Component Parameters 498
8.5.9 Managing MRS Role Instances 502
8.5.10 Managing MRS Role Instance Groups 503
8.5.11 Modifying MRS Role Instance Parameters 505
8.5.12 Perform an Active/Standby Switchover for MRS Role Instances 507
8.5.13 Decommissioning and Recommissioning an MRS Role Instance 508
8.5.14 Enabling and Disabling Ranger Authentication for an MRS Component 511
8.5.15 Accessing Web UIs of Open-Source Components Managed in MRS Clusters 513

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.6 Managing MRS Cluster Nodes 515
8.6.1 Checking the Running Status of an MRS Cluster Node 515
8.6.2 Starting and Stopping All Roles on an MRS Cluster Node 518
8.6.3 Isolating an MRS Cluster Node 519
8.6.4 Modifying the Rack Information of an MRS Cluster Node 521
8.6.5 Managing MRS Cluster Node Groups 524
8.6.6 Adding a Tag to an MRS Cluster/Node 526
8.6.7 Configuring Bootstrap Actions for an MRS Cluster Node 530
8.6.7.1 MRS Bootstrap Action Overview 530
8.6.7.2 Preparing the Bootstrap Action Script for an MRS Node 531
8.6.7.3 Adding MRS Node Bootstrap Actions and Installing Third-Party Software 532
8.6.7.4 Viewing the Bootstrap Action Execution Records of an MRS Node 536
8.7 Managing the MRS Cluster Client 537
8.7.1 Updating the MRS Cluster Client After the Server Configuration Expires 537
8.7.2 Viewing the Installed MRS Cluster Client 543
8.7.3 Batch Upgrading MRS Cluster Clients 543
8.8 Managing MRS Cluster Jobs 546
8.8.1 Stopping and Deleting an MRS Cluster Job 546
8.8.2 Configuring Notification Rules for MRS Jobs 547
8.9 Managing MRS Cluster Tenants 550
8.9.1 Introduction to MRS Multi-Tenancy 550
8.9.2 Using MRS Multi-Tenancy 561
8.9.3 Configuring MRS Tenants 563
8.9.3.1 Creating an MRS Tenant 564
8.9.3.2 Creating an MRS Sub-Tenant 578
8.9.3.3 Binding a Tenant to an MRS Cluster User 596
8.9.3.4 Adding an MRS Tenant Resource Pool 598
8.9.3.5 Configuring the Queue Capacity Policy of a Resource Pool 600
8.9.3.6 Configuring the MRS Tenant Queue 603
8.9.4 Managing MRS Tenant Resources 612
8.9.4.1 Managing the MRS Tenant Resource Directory 612
8.9.4.2 Managing MRS Tenant Resource Pools 614
8.9.4.3 Clearing the MRS Tenant Queue Configuration 616
8.9.4.4 Restoring MRS Tenant Data After YARN Is Reinstalled 617
8.9.4.5 Deleting an MRS Tenant 618
8.9.4.6 Managing Global User Policies When Using Superior Scheduler 620
8.9.4.7 Clearing Tenant's Non-Associated Queues Using Capacity Scheduler 621
8.9.5 Switching the MRS Tenant Resource Scheduler 622
8.10 Managing MRS Cluster Users 626
8.10.1 Cluster User Permissions 626
8.10.1.1 MRS Cluster User Permission Model 626
8.10.1.2 MRS Cluster User Identity Authentication Policy 628

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.1.3 MRS Cluster User Permission Authentication Policy....630
8.10.1.4 Default Permissions of the MRS Cluster....633
8.10.1.5 Synchronizing IAM Users to MRS....636
8.10.2 MRS Cluster User Accounts....641
8.10.3 Managing MRS Cluster Roles....695
8.10.4 Managing MRS Cluster User Groups....698
8.10.5 Managing MRS Cluster Users....700
8.10.5.1 Creating an MRS Cluster User....700
8.10.5.2 Modifying MRS Cluster User Information....703
8.10.5.3 Locking an MRS Cluster User....704
8.10.5.4 Deleting an MRS Cluster User....705
8.10.5.5 Initializing MRS Cluster User Passwords....706
8.10.5.6 Downloading MRS Cluster User Credentials....707
8.10.6 Unlocking an MRS Cluster User....710
8.10.6.1 Unlocking an LDAP User in the MRS Cluster....710
8.10.6.2 Unlocking the LDAP Management Account of the MRS Cluster....712
8.10.7 Configuring Password Policies for MRS Cluster Users....713
8.10.8 Configuring the Private Attribute of MRS Cluster Users....717
8.11 Managing MRS Cluster Metadata....719
8.11.1 MRS Cluster Metadata Storage in an External Data Source....719
8.11.2 Storing Ranger Metadata to RDS....722
8.11.3 Storing Hive Metadata to RDS....728
8.11.4 Managing MRS Cluster Data Connections....732
8.12 Managing Static Service Resources in an MRS Cluster....733
8.12.1 Overview of Static Service Resources....733
8.12.2 Configuring Static Resources for an MRS Cluster....734
8.12.3 Checking the Static Resources of an MRS Cluster....739
8.13 Managing SQL Inspection Rules for an MRS Cluster....741
8.13.1 SQL Inspector Overview....741
8.13.2 Adding an MRS SQL Inspection Rule....742
8.13.3 Configuring Hive SQL Inspection Rules....749
8.13.4 Configuring ClickHouse SQL Inspection Rules....751
8.13.5 Configuring HetuEngine SQL Inspection Rules....753
8.13.6 Configuring Spark SQL Inspection Rules....754
8.13.7 Configuring Doris SQL Inspection....757
9 MRS Cluster O&M....760
9.1 Cluster O&M....760
9.2 Logging In to an MRS Cluster....764
9.2.1 Checking MRS Active/Standby Management Nodes....764
9.2.2 Logging In to an MRS Cluster Node....765
9.3 Viewing MRS Cluster Monitoring Metrics....769
9.3.1 Viewing MRS Cluster Resource Monitoring Metrics....769

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.3.2 Viewing MRS Cluster Component Monitoring Metrics....773
9.3.3 Viewing MRS Node Resource Monitoring Metrics....778
9.3.4 Dumping MRS Cluster Monitoring Data....782
9.4 Checking MRS Cluster Health....786
9.4.1 Performing a Health Check for an MRS Cluster....786
9.4.2 Performing Health Checks on MRS Cluster Nodes....788
9.4.3 Viewing and Exporting a Health Check Report....788
9.5 Adjusting the Capacity of an MRS Cluster....790
9.5.1 Scaling Out an MRS Cluster....790
9.5.2 Scaling In an MRS Cluster....794
9.5.3 Scaling In ClickHouseServer Nodes....798
9.5.4 MRS Task Node Auto Scaling....806
9.5.4.1 Automatic Scaling of Task Nodes in an MRS Cluster....806
9.5.4.2 Adding an Auto Scaling Policy for MRS Task Nodes....813
9.5.4.3 Managing MRS Cluster Auto Scaling Policies....824
9.6 MRS Cluster Data Backup and Restoration....825
9.6.1 Backing Up and Restoring MRS Cluster Data....826
9.6.2 Enabling MRS Inter-Cluster Replication....833
9.6.3 Creating an MRS Cluster Data Backup Task....834
9.6.4 Creating an MRS Cluster Data Restoration Task....836
9.6.5 Backing Up MRS Cluster Component Data....837
9.6.5.1 Backing Up Manager Data (MRS 2.x and Earlier)....837
9.6.5.2 Backing Up Manager Data (MRS 3.x and Later Versions)....839
9.6.5.3 Backing Up CDL Service Data....844
9.6.5.4 Backing Up ClickHouse Metadata....846
9.6.5.5 Backing Up ClickHouse Service Data....851
9.6.5.6 Backing Up DBService Data....857
9.6.5.7 Backing Up Doris Data....862
9.6.5.8 Backing Up Flink Metadata....866
9.6.5.9 Backing Up HBase Metadata....869
9.6.5.10 Backing Up HBase Service Data....874
9.6.5.11 Backing Up HDFS NameNode Data....880
9.6.5.12 Backing Up HDFS Service Data....885
9.6.5.13 Backing Up Hive Service Data....892
9.6.5.14 Backing Up IoTDB Metadata....900
9.6.5.15 Backing Up IoTDB Service Data....903
9.6.5.16 Backing Up Kafka Metadata....906
9.6.6 Restoring MRS Cluster Component Data....911
9.6.6.1 Restoring Manager Data (MRS 2.x and Earlier)....911
9.6.6.2 Restoring Manager Data (MRS 3.x and Later Versions)....914
9.6.6.3 Restoring CDL Service Data....919
9.6.6.4 Restoring ClickHouse Metadata....921

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.6.6.5 Restoring ClickHouse Service Data....924
9.6.6.6 Restoring DBService Metadata....928
9.6.6.7 Restoring Doris Service Data....932
9.6.6.8 Restoring Flink Metadata....935
9.6.6.9 Restoring HBase Metadata....939
9.6.6.10 Restoring HBase Service Data....943
9.6.6.11 Restoring HDFS NameNode Metadata....948
9.6.6.12 Restoring HDFS Service Data....952
9.6.6.13 Restoring Hive Service Data....958
9.6.6.14 Restoring IoTDB Metadata....964
9.6.6.15 Restoring IoTDB Service Data....967
9.6.6.16 Restoring Kafka Metadata....970
9.6.7 Managing MRS Cluster Backup and Restoration Tasks....974
9.6.8 Using HDFS Snapshots to Quickly Restore Component Service Data....977
9.7 MRS Cluster Patching....978
9.7.1 Viewing Patch Information for an MRS Cluster....978
9.7.2 Patching an MRS Cluster....979
9.7.3 Applying Rolling Patches for an MRS Cluster....980
9.7.4 Patching Hosts Isolated in an MRS Cluster....983
9.8 Viewing Logs of an MRS Cluster....984
9.8.1 Overview of MRS Cluster Logs....984
9.8.2 Viewing MRS Operation Logs....1013
9.8.3 Viewing MRS Cluster History....1014
9.8.4 Viewing MRS Cluster Audit Logs....1016
9.8.5 Viewing Role Instance Logs of MRS Components....1019
9.8.6 Searching for MRS Cluster Logs Online....1020
9.8.7 Downloading MRS Cluster Logs....1023
9.8.8 Collecting MRS Cluster Service Stack Information....1024
9.8.9 Configuring Default Log Level and Archive File Size for MRS Components....1025
9.8.10 Configuring the Number of Local Backups of MRS Cluster Audit Logs....1027
9.8.11 Configuring Dumping for MRS Cluster Audit Logs....1027
9.9 MRS Cluster Security Configuration....1031
9.9.1 Cluster Mutual Trust Management....1031
9.9.1.1 Overview of Mutual Trust Between MRS Clusters....1031
9.9.1.2 Changing the System Domain Name of an MRS Cluster....1031
9.9.1.3 Configuring Mutual Trust Between MRS Clusters....1035
9.9.1.4 Configuring User Permissions for Mutually Trusted MRS Clusters....1042
9.9.2 MRS Cluster Security Hardening....1043
9.9.2.1 MRS Cluster Security Hardening Policies....1043
9.9.2.2 Configuring Hadoop Data Encryption During Transmission....1044
9.9.2.3 Configuring Kafka Data Encryption During Transmission....1048
9.9.2.4 Encrypting Data Transmission Between the Controller and Agent....1049

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.9.2.5 Configuring a Trusted IP Address to Access LDAP....1050
9.9.2.6 Configuring the IP Address Whitelist for Modifying Data in an HBase Read-Only Cluster....1053
9.9.2.7 Configuring LDAP Output Audit Logs....1054
9.9.2.8 Updating Encryption Keys of an MRS Cluster....1055
9.9.2.9 Updating the SSH Key of User omm on MRS Cluster Nodes....1057
9.9.2.10 Enabling and Disabling Permission Verification on MRS Cluster Components....1059
9.9.2.11 Allowing External Users to Access MRS Clusters in Normal Mode....1060
9.9.2.12 Modifying the Session Timeout Duration for MRS Cluster Management Pages....1062
9.9.2.13 Configuring Secure Communication Authorization for an MRS Cluster....1063
9.9.3 Changing the Passwords for System Users of an MRS Cluster....1065
9.9.3.1 Changing or Resetting the Password for User admin of an MRS Cluster....1065
9.9.3.2 Changing the Passwords for OS Users of an MRS Cluster Node....1068
9.9.3.3 Changing the Password for the Kerberos Administrator of an MRS Cluster....1069
9.9.3.4 Changing the Passwords for Manager Users of an MRS Cluster....1071
9.9.3.5 Changing the Password for a Regular LDAP User of an MRS Cluster....1073
9.9.3.6 Changing the LDAP Administrator Password for an MRS Cluster....1076
9.9.3.7 Changing the Passwords for MRS Cluster Component Running Users....1077
9.9.4 Changing the Passwords for Database Users of an MRS Cluster....1080
9.9.4.1 Changing the Password for the OMS Database Administrator....1080
9.9.4.2 Changing the Password for an OMS Database Access User....1080
9.9.4.3 Changing the Passwords for Database Users of MRS Cluster Components....1082
9.9.4.4 Resetting the MRS Component Database User Password....1083
9.9.4.5 Resetting the Password for User omm in DBService....1084
9.9.4.6 Changing the Password for User compdbuser of the DBService Database....1085
9.10 Viewing and Configuring MRS Alarm Events....1086
9.10.1 Viewing MRS Cluster Events....1086
9.10.2 Viewing Alarms of an MRS Cluster....1090
9.10.3 Configuring Alarm Thresholds for an MRS Cluster....1095
9.10.4 Configuring Alarm Masking for an MRS Cluster....1116
9.10.5 Connecting an MRS Cluster to SNMP to Report Alarms....1118
9.10.6 Connecting an MRS Cluster to the Syslog Server to Report Alarms....1120
9.10.7 Periodically Backing Up Alarm and Audit Information....1126
9.10.8 Enabling the MRS Cluster Maintenance Mode to Disable Alarm Reporting....1127
9.10.9 Configuring Notifications for MRS Cluster Alarms and Events....1129
9.11 MRS Cluster Alarm Handling Reference....1133
9.11.1 ALM-12001 Audit Log Dumping Failure....1133
9.11.2 ALM-12004 OLdap Resource Abnormal....1136
9.11.3 ALM-12005 OKerberos Resource Abnormal....1138
9.11.4 ALM-12006 Node Fault....1139
9.11.5 ALM-12007 Process Fault....1144
9.11.6 ALM-12010 Manager Heartbeat Interruption Between the Active and Standby Nodes....1149
9.11.7 ALM-12011 Manager Data Synchronization Exception Between the Active and Standby Nodes.1152

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.8 ALM-12012 NTP Service Is Abnormal 1155
9.11.9 ALM-12014 Partition Lost 1162
9.11.10 ALM-12015 Partition Filesystem Readonly 1165
9.11.11 ALM-12016 CPU Usage Exceeds the Threshold 1166
9.11.12 ALM-12017 Insufficient Disk Capacity 1168
9.11.13 ALM-12018 Memory Usage Exceeds the Threshold 1171
9.11.14 ALM-12027 Host PID Usage Exceeds the Threshold 1173
9.11.15 ALM-12028 Number of Processes in the D State and Z State on a Host Exceeds the Threshold 1175
9.11.16 ALM-12033 Slow Disk Fault 1177
9.11.17 ALM-12034 Periodical Backup Failure 1184
9.11.18 ALM-12035 Unknown Data Status After Recovery Task Failure 1187
9.11.19 ALM-12037 NTP Server Abnormal 1188
9.11.20 ALM-12038 Monitoring Indicator Dumping Failure 1191
9.11.21 ALM-12039 Active/Standby OMS Databases Not Synchronized 1194
9.11.22 ALM-12040 Insufficient System Entropy 1196
9.11.23 ALM-12041 Incorrect Permission on Key Files 1199
9.11.24 ALM-12042 Incorrect Configuration of Key Files 1201
9.11.25 ALM-12045 Read Packet Dropped Rate Exceeds the Threshold 1204
9.11.26 ALM-12046 Write Packet Dropped Rate Exceeds the Threshold 1207
9.11.27 ALM-12047 Read Packet Error Rate Exceeds the Threshold 1209
9.11.28 ALM-12048 Write Packet Error Rate Exceeds the Threshold 1211
9.11.29 ALM-12049 Network Read Throughput Rate Exceeds the Threshold 1213
9.11.30 ALM-12050 Network Write Throughput Rate Exceeds the Threshold 1216
9.11.31 ALM-12051 Disk Inode Usage Exceeds the Threshold 1218
9.11.32 ALM-12052 TCP Temporary Port Usage Exceeds the Threshold 1220
9.11.33 ALM-12053 Host File Handle Usage Exceeds the Threshold 1223
9.11.34 ALM-12054 Invalid Certificate File 1226
9.11.35 ALM-12055 Certificate File Is About to Expire 1229
9.11.36 ALM-12057 Metadata Not Configured with the Task to Periodically Back Up Data to a Third-Party Server 1232
9.11.37 ALM-12061 Process Usage Exceeds the Threshold 1233
9.11.38 ALM-12062 OMS Parameter Configurations Mismatch with the Cluster Scale 1236
9.11.39 ALM-12063 Unavailable Disk 1239
9.11.40 ALM-12064 Host Random Port Range Conflicts with Cluster Used Port 1241
9.11.41 ALM-12066 Trust Relationships Between Nodes Become Invalid 1242
9.11.42 ALM-12067 Tomcat Resource Is Abnormal 1246
9.11.43 ALM-12068 ACS Resource Exception 1248
9.11.44 ALM-12069 AOS Resource Exception 1250
9.11.45 ALM-12070 Controller Resource Is Abnormal 1251
9.11.46 ALM-12071 Httpd Resource Is Abnormal 1253
9.11.47 ALM-12072 FloatIP Resource Is Abnormal 1255
9.11.48 ALM-12073 CEP Resource Is Abnormal 1257
9.11.49 ALM-12074 FMS Resource Is Abnormal 1259

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.50 ALM-12075 PMS Resource Is Abnormal 1261
9.11.51 ALM-12076 GaussDB Resource Is Abnormal 1263
9.11.52 ALM-12077 User omm Expired 1265
9.11.53 ALM-12078 Password of User omm Expired 1267
9.11.54 ALM-12079 User omm Is About to Expire 1269
9.11.55 ALM-12080 Password of User omm Is About to Expire 1271
9.11.56 ALM-12081 User ommdba Expired 1272
9.11.57 ALM-12082 User ommdba Is About to Expire 1274
9.11.58 ALM-12083 Password of User ommdba Is About to Expire 1276
9.11.59 ALM-12084 Password of User ommdba Expired 1277
9.11.60 ALM-12085 Service Audit Log Dump Failure 1279
9.11.61 ALM-12087 System Is in the Upgrade Observation Period 1282
9.11.62 ALM-12089 Inter-Node Network Is Abnormal 1284
9.11.63 ALM-12091 Abnormal disaster Resources 1286
9.11.64 ALM-12101 AZ Unhealthy 1288
9.11.65 ALM-12102 AZ HA Component Is Not Deployed Based on DR Requirements 1290
9.11.66 ALM-12103 Executor Resource Exception 1291
9.11.67 ALM-12104 Abnormal Knox Resources 1293
9.11.68 ALM-12110 Failed to get ECS temporary AK/SK 1294
9.11.69 ALM-12172 Failed to Report Metrics to Cloud Eye 1296
9.11.70 ALM-12180 Suspended Disk I/O 1298
9.11.71 ALM-12186 CGroup Task Usage Exceeds the Threshold 1302
9.11.72 ALM-12187 Failed to Expand Disk Partition Capacity 1304
9.11.73 ALM-12188 diskmgt Disk Monitoring Unavailable 1306
9.11.74 ALM-12190 Number of Knox Connections Exceeds the Threshold 1308
9.11.75 ALM-12191 Disk I/O Usage Exceeds the Threshold 1310
9.11.76 ALM-12192 Host Load Exceeds the Threshold 1312
9.11.77 ALM-12200 Password Is About to Expire 1315
9.11.78 ALM-12201 Process CPU Usage Exceeds the Threshold 1316
9.11.79 ALM-12202 Process Memory Usage Exceeds the Threshold 1319
9.11.80 ALM-12203 Process Full GC Duration Exceeds the Threshold 1321
9.11.81 ALM-12204 Wait Duration of a Disk Read Exceeds the Threshold 1322
9.11.82 ALM-12205 Wait Duration of a Disk Write Exceeds the Threshold 1325
9.11.83 ALM-12206 Password Has Expired 1327
9.11.84 ALM-12207 Slow Disk Processing Timeout 1329
9.11.85 ALM-13000 ZooKeeper Service Unavailable 1331
9.11.86 ALM-13001 Available ZooKeeper Connections Are Insufficient 1334
9.11.87 ALM-13002 ZooKeeper Direct Memory Usage Exceeds the Threshold 1337
9.11.88 ALM-13003 GC Duration of the ZooKeeper Process Exceeds the Threshold 1339
9.11.89 ALM-13004 ZooKeeper Heap Memory Usage Exceeds the Threshold 1342
9.11.90 ALM-13005 Failed to Set the Quota of Top Directories of ZooKeeper Components 1344
9.11.91 ALM-13006 Znode Number or Capacity Exceeds the Threshold 1346

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.92 ALM-13007 Available ZooKeeper Client Connections Are Insufficient 1348
9.11.93 ALM-13008 ZooKeeper Znode Usage Exceeds the Threshold 1350
9.11.94 ALM-13009 ZooKeeper ZNode Capacity Usage Exceeds the Threshold 1352
9.11.95 ALM-13010 Znode Usage of a Directory with Quota Configured Exceeds the Threshold 1354
9.11.96 ALM-14000 HDFS Service Unavailable 1356
9.11.97 ALM-14001 HDFS Disk Usage Exceeds the Threshold 1359
9.11.98 ALM-14002 DataNode Disk Usage Exceeds the Threshold 1362
9.11.99 ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold 1365
9.11.100 ALM-14006 Number of HDFS Files Exceeds the Threshold 1368
9.11.101 ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold 1371
9.11.102 ALM-14008 DataNode Heap Memory Usage Exceeds the Threshold 1374
9.11.103 ALM-14009 Number of Dead DataNodes Exceeds the Threshold 1377
9.11.104 ALM-14010 NameService Service Is Abnormal 1380
9.11.105 ALM-14011 DataNode Data Directory Is Not Configured Properly 1384
9.11.106 ALM-14012 JournalNode Is Out of Synchronization 1388
9.11.107 ALM-14013 Failed to Update the NameNode FsImage File 1390
9.11.108 ALM-14014 NameNode GC Time Exceeds the Threshold 1395
9.11.109 ALM-14015 DataNode GC Time Exceeds the Threshold 1398
9.11.110 ALM-14016 DataNode Direct Memory Usage Exceeds the Threshold 1400
9.11.111 ALM-14017 NameNode Direct Memory Usage Exceeds the Threshold 1402
9.11.112 ALM-14018 NameNode Non-heap Memory Usage Exceeds the Threshold 1405
9.11.113 ALM-14019 DataNode Non-heap Memory Usage Exceeds the Threshold 1408
9.11.114 ALM-14020 Number of Entries in the HDFS Directory Exceeds the Threshold 1410
9.11.115 ALM-14021 NameNode Average RPC Processing Time Exceeds the Threshold 1413
9.11.116 ALM-14022 NameNode Average RPC Queuing Time Exceeds the Threshold 1416
9.11.117 ALM-14023 Percentage of Total Reserved Disk Space for Replicas Exceeds the Threshold 1420
9.11.118 ALM-14024 Tenant Space Usage Exceeds the Threshold 1423
9.11.119 ALM-14025 Tenant File Object Usage Exceeds the Threshold 1425
9.11.120 ALM-14026 Blocks on DataNode Exceed the Threshold 1427
9.11.121 ALM-14027 DataNode Disk Fault 1431
9.11.122 ALM-14028 Number of Blocks to Be Supplemented Exceeds the Threshold 1433
9.11.123 ALM-14029 Number of Blocks in a Replica Exceeds the Threshold 1436
9.11.124 ALM-14030 HDFS Allows Write of Single-Replica Data 1438
9.11.125 ALM-14031 DataNode Process Is Abnormal 1439
9.11.126 ALM-14032 JournalNode Process Is Abnormal 1441
9.11.127 ALM-14033 ZKFC Process Is Abnormal 1443
9.11.128 ALM-14034 Router Process Is Abnormal 1445
9.11.129 ALM-14035 HttpFS Process Is Abnormal 1447
9.11.130 ALM-14036 NameNode Is In Safe Mode 1449
9.11.131 ALM-14037 DataNodes Outside the Cluster 1451
9.11.132 ALM-14038 Router Heap Memory Usage Exceeds the Threshold 1453
9.11.133 ALM-14039 Slow DataNodes Exist in the Cluster 1455

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.134 ALM-16000 Percentage of Sessions Connected to the HiveServer to Maximum Number Allowed Exceeds the Threshold 1458
9.11.135 ALM-16001 Hive Warehouse Space Usage Exceeds the Threshold 1460
9.11.136 ALM-16002 Hive SQL Execution Success Rate Is Lower Than the Threshold 1462
9.11.137 ALM-16003 Background Thread Usage Exceeds the Threshold 1465
9.11.138 ALM-16004 Hive Service Unavailable 1467
9.11.139 ALM-16005 The Heap Memory Usage of the Hive Process Exceeds the Threshold 1472
9.11.140 ALM-16006 The Direct Memory Usage of the Hive Process Exceeds the Threshold 1474
9.11.141 ALM-16007 Hive GC Time Exceeds the Threshold 1477
9.11.142 ALM-16008 Non-Heap Memory Usage of the Hive Process Exceeds the Threshold 1480
9.11.143 ALM-16009 Map Number Exceeds the Threshold 1482
9.11.144 ALM-16045 Hive Data Warehouse Is Deleted 1484
9.11.145 ALM-16046 Hive Data Warehouse Permission Is Modified 1485
9.11.146 ALM-16047 HiveServer Has Been Deregistered from ZooKeeper 1487
9.11.147 ALM-16048 Tez or Spark Library Path Does Not Exist 1490
9.11.148 ALM-17003 Oozie Service Unavailable 1491
9.11.149 ALM-17004 Oozie Heap Memory Usage Exceeds the Threshold 1495
9.11.150 ALM-17005 Oozie Non Heap Memory Usage Exceeds the Threshold 1497
9.11.151 ALM-17006 Oozie Direct Memory Usage Exceeds the Threshold 1500
9.11.152 ALM-17007 Garbage Collection (GC) Time of the Oozie Process Exceeds the Threshold 1502
9.11.153 ALM-17008 Abnormal Connection Between Oozie and ZooKeeper 1504
9.11.154 ALM-17009 Abnormal Connection Between Oozie and DBService 1506
9.11.155 ALM-17010 Abnormal Connection Between Oozie and HDFS 1508
9.11.156 ALM-17011 Abnormal Connection Between Oozie and YARN 1510
9.11.157 ALM-18000 Yarn Service Unavailable 1512
9.11.158 ALM-18002 NodeManager Heartbeat Lost 1514
9.11.159 ALM-18003 NodeManager Unhealthy 1517
9.11.160 ALM-18008 Heap Memory Usage of ResourceManager Exceeds the Threshold 1520
9.11.161 ALM-18009 Heap Memory Usage of JobHistoryServer Exceeds the Threshold 1523
9.11.162 ALM-18010 ResourceManager GC Time Exceeds the Threshold 1525
9.11.163 ALM-18011 NodeManager GC Time Exceeds the Threshold 1528
9.11.164 ALM-18012 JobHistoryServer GC Time Exceeds the Threshold 1530
9.11.165 ALM-18013 ResourceManager Direct Memory Usage Exceeds the Threshold 1532
9.11.166 ALM-18014 NodeManager Direct Memory Usage Exceeds the Threshold 1535
9.11.167 ALM-18015 JobHistoryServer Direct Memory Usage Exceeds the Threshold 1537
9.11.168 ALM-18016 Non Heap Memory Usage of ResourceManager Exceeds the Threshold 1539
9.11.169 ALM-18017 Non Heap Memory Usage of NodeManager Exceeds the Threshold 1542
9.11.170 ALM-18018 NodeManager Heap Memory Usage Exceeds the Threshold 1544
9.11.171 ALM-18019 Non Heap Memory Usage of JobHistoryServer Exceeds the Threshold 1547
9.11.172 ALM-18020 Yarn Task Execution Timeout 1549
9.11.173 ALM-18021 MapReduce Service Unavailable 1552
9.11.174 ALM-18022 Insufficient Yarn Queue Resources 1554
9.11.175 ALM-18023 Number of Pending Yarn Tasks Exceeds the Threshold 1557

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.176 ALM-18024 Pending Yarn Memory Usage Exceeds the Threshold 1559
9.11.177 ALM-18025 Number of Terminated Yarn Tasks Exceeds the Threshold 1561
9.11.178 ALM-18026 Number of Failed Yarn Tasks Exceeds the Threshold 1563
9.11.179 ALM-18027 JobHistoryServer Process Is Abnormal 1565
9.11.180 ALM-18028 TimeLineServer Process Is Abnormal 1567
9.11.181 ALM-19000 HBase Service Unavailable 1569
9.11.182 ALM-19006 HBase Replication Sync Failed 1575
9.11.183 ALM-19007 HBase GC Time Exceeds the Threshold 1579
9.11.184 ALM-19008 Heap Memory Usage of the HBase Process Exceeds the Threshold 1581
9.11.185 ALM-19009 Direct Memory Usage of the HBase Process Exceeds the Threshold 1584
9.11.186 ALM-19011 RegionServer Region Number Exceeds the Threshold 1586
9.11.187 ALM-19012 HBase System Table Directory or File Lost 1590
9.11.188 ALM-19013 Duration of Regions in transaction State Exceeds the Threshold 1592
9.11.189 ALM-19014 Capacity Quota Usage on ZooKeeper Exceeds the Threshold Severely 1594
9.11.190 ALM-19015 Quantity Quota Usage on ZooKeeper Exceeds the Threshold 1597
9.11.191 ALM-19016 Quantity Quota Usage on ZooKeeper Exceeds the Threshold Severely 1599
9.11.192 ALM-19017 Capacity Quota Usage on ZooKeeper Exceeds the Threshold 1602
9.11.193 ALM-19018 HBase Compaction Queue Size Exceeds the Threshold 1604
9.11.194 ALM-19019 Number of HBase HFiles to Be Synchronized Exceeds the Threshold 1606
9.11.195 ALM-19020 Number of HBase WAL Files to Be Synchronized Exceeds the Threshold 1609
9.11.196 ALM-19021 Handler Usage of RegionServer Exceeds the Threshold 1612
9.11.197 ALM-19022 HBase Hotspot Detection Is Unavailable 1615
9.11.198 ALM-19023 Region Traffic Restriction for HBase 1618
9.11.199 ALM-19024 RPC Requests P99 Latency on RegionServer Exceeds the Threshold 1620
9.11.200 ALM-19025 Damaged StoreFile in HBase 1623
9.11.201 ALM-19026 Damaged WAL Files in HBase 1625
9.11.202 ALM-20002 Hue Service Unavailable 1627
9.11.203 ALM-23001 Loader Service Unavailable 1629
9.11.204 ALM-23003 Loader Task Execution Failure 1633
9.11.205 ALM-23004 Loader Heap Memory Usage Exceeds the Threshold 1636
9.11.206 ALM-23005 Loader Non-Heap Memory Usage Exceeds the Threshold 1638
9.11.207 ALM-23006 Loader Direct Memory Usage Exceeds the Threshold 1640
9.11.208 ALM-23007 Garbage Collection (GC) Time of the Loader Process Exceeds the Threshold 1642
9.11.209 ALM-24000 Flume Service Unavailable 1644
9.11.210 ALM-24001 Flume Agent Exception 1646
9.11.211 ALM-24003 Flume Client Connection Interrupted 1649
9.11.212 ALM-24004 Exception Occurs When Flume Reads Data 1652
9.11.213 ALM-24005 Exception Occurs When Flume Transmits Data 1654
9.11.214 ALM-24006 Heap Memory Usage of Flume Server Exceeds the Threshold 1657
9.11.215 ALM-24007 Flume Server Direct Memory Usage Exceeds the Threshold 1659
9.11.216 ALM-24008 Flume Server Non-Heap Memory Usage Exceeds the Threshold 1662
9.11.217 ALM-24009 Flume Server Garbage Collection (GC) Time Exceeds the Threshold 1664

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.218 ALM-24010 Flume Certificate File Is Invalid or Damaged 1666
9.11.219 ALM-24011 Flume Certificate File Is About to Expire 1668
9.11.220 ALM-24012 Flume Certificate File Has Expired 1670
9.11.221 ALM-24013 Flume MonitorServer Certificate File Is Invalid or Damaged 1673
9.11.222 ALM-24014 Flume MonitorServer Certificate Is About to Expire 1675
9.11.223 ALM-24015 Flume MonitorServer Certificate File Has Expired 1677
9.11.224 ALM-25000 LdapServer Service Unavailable 1680
9.11.225 ALM-25004 Abnormal LdapServer Data Synchronization 1682
9.11.226 ALM-25005 nscd Service Exception 1685
9.11.227 ALM-25006 Sssd Service Exception 1689
9.11.228 ALM-25007 Number of SlapdServer Connections Exceeds the Threshold 1692
9.11.229 ALM-25008 SlapdServer CPU Usage Exceeds the Threshold 1694
9.11.230 ALM-25500 KrbServer Service Unavailable 1697
9.11.231 ALM-25501 Too Many KerberosServer Requests 1699
9.11.232 ALM-26051 Storm Service Unavailable 1701
9.11.233 ALM-26052 Number of Available Supervisors of the Storm Service Is Less Than the Threshold 1703
9.11.234 ALM-26053 Storm Slot Usage Exceeds the Threshold 1705
9.11.235 ALM-26054 Nimbus Heap Memory Usage Exceeds the Threshold 1707
9.11.236 ALM-27001 DBService Service Unavailable 1710
9.11.237 ALM-27003 DBService Heartbeat Interruption Between the Active and Standby Nodes 1713
9.11.238 ALM-27004 Data Inconsistency Between Active and Standby DBServices 1715
9.11.239 ALM-27005 Database Connections Usage Exceeds the Threshold 1717
9.11.240 ALM-27006 Disk Space Usage of the Data Directory Exceeds the Threshold 1720
9.11.241 ALM-27007 Database Enters the Read-Only Mode 1723
9.11.242 ALM-29000 Impala Service Unavailable 1725
9.11.243 ALM-29004 Impalad Process Memory Usage Exceeds the Threshold 1728
9.11.244 ALM-29005 Number of JDBC Connections to Impalad Exceeds the Threshold 1730
9.11.245 ALM-29006 Number of ODBC Connections to Impalad Exceeds the Threshold 1732
9.11.246 ALM-29007 Impalad Process Memory Usage Exceeds the Threshold 1735
9.11.247 ALM-29008 Number of ODBC Connections to Impalad Exceeds the Threshold 1737
9.11.248 ALM-29010 Number of Queries Being Submitted by Impalad Exceeds the Threshold 1739
9.11.249 ALM-29011 Number of Queries Being Executed by Impalad Exceeds the Threshold 1741
9.11.250 ALM-29012 Number of Queries Being Waited by Impalad Exceeds the Threshold 1743
9.11.251 ALM-29013 Impalad FGC Time Exceeds the Threshold 1745
9.11.252 ALM-29014 Catalog FGC Time Exceeds the Threshold 1747
9.11.253 ALM-29015 Catalog Process Memory Usage Exceeds the Threshold 1749
9.11.254 ALM-29016 Impalad Instance in the Sub-healthy State 1751
9.11.255 ALM-29104 Tserver Process Memory Usage Exceeds the Threshold 1753
9.11.256 ALM-29106 Tserver Process CPU Usage Exceeds the Threshold 1754
9.11.257 ALM-29107 Tserver Process Memory Usage Exceeds the Threshold 1756
9.11.258 ALM-38000 Kafka Service Unavailable 1758
9.11.259 ALM-38001 Insufficient Kafka Disk Capacity 1760

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.260 ALM-38002 Kafka Heap Memory Usage Exceeds the Threshold....1765
9.11.261 ALM-38004 Kafka Direct Memory Usage Exceeds the Threshold....1767
9.11.262 ALM-38005 GC Duration of the Broker Process Exceeds the Threshold....1770
9.11.263 ALM-38006 Percentage of Kafka Partitions That Are Not Completely Synchronized Exceeds the Threshold....1772
9.11.264 ALM-38007 Status of Kafka Default User Is Abnormal....1774
9.11.265 ALM-38008 Abnormal Kafka Data Directory Status....1776
9.11.266 ALM-38009 Busy Broker Disk I/Os (Applicable to Versions Later Than MRS 3.1.0)....1778
9.11.267 ALM-38010 Topics with Single Replica....1781
9.11.268 ALM-38011 User Connection Usage on Broker Exceeds the Threshold....1783
9.11.269 ALM-38012 Number of Broker Partitions Exceeds the Threshold....1787
9.11.270 ALM-38013 Produce Request Latency in the Request Queue Exceeds the Threshold....1789
9.11.271 ALM-38014 Total Produce Request Latency Exceeds the Threshold....1792
9.11.272 ALM-38015 Fetch Request Latency in the Request Queue Exceeds the Threshold....1795
9.11.273 ALM-38016 Total Fetch Request Latency Exceeds the Threshold....1798
9.11.274 ALM-38017 Partition Reassignment Duration Exceeds the Threshold....1800
9.11.275 ALM-38018 Kafka Consumer Lag....1802
9.11.276 ALM-43001 Spark2x Service Unavailable....1805
9.11.277 ALM-43006 Heap Memory Usage of the JobHistory2x Process Exceeds the Threshold....1808
9.11.278 ALM-43007 Non-Heap Memory Usage of the JobHistory2x Process Exceeds the Threshold....1811
9.11.279 ALM-43008 The Direct Memory Usage of the JobHistory2x Process Exceeds the Threshold....1813
9.11.280 ALM-43009 JobHistory2x Process GC Time Exceeds the Threshold....1816
9.11.281 ALM-43010 Heap Memory Usage of the JDBCServer2x Process Exceeds the Threshold....1818
9.11.282 ALM-43011 Non-Heap Memory Usage of the JDBCServer2x Process Exceeds the Threshold....1820
9.11.283 ALM-43012 Direct Heap Memory Usage of the JDBCServer2x Process Exceeds the Threshold1823
9.11.284 ALM-43013 JDBCServer2x Process GC Time Exceeds the Threshold....1825
9.11.285 ALM-43017 JDBCServer2x Process Full GC Number Exceeds the Threshold....1827
9.11.286 ALM-43018 JobHistory2x Process Full GC Number Exceeds the Threshold....1830
9.11.287 ALM-43019 Heap Memory Usage of the IndexServer2x Process Exceeds the Threshold....1832
9.11.288 ALM-43020 Non-Heap Memory Usage of the IndexServer2x Process Exceeds the Threshold....1834
9.11.289 ALM-43021 Direct Memory Usage of the IndexServer2x Process Exceeds the Threshold....1837
9.11.290 ALM-43022 IndexServer2x Process GC Time Exceeds the Threshold....1839
9.11.291 ALM-43023 IndexServer2x Process Full GC Number Exceeds the Threshold....1841
9.11.292 ALM-43028 JDBCServer Session Overflow....1843
9.11.293 ALM-43029 JDBCServer Job Submission Timed Out....1845
9.11.294 ALM-44004 Presto Coordinator Resource Group Queuing Tasks Exceed the Threshold....1847
9.11.295 ALM-44005 Presto Coordinator Process GC Time Exceeds the Threshold....1849
9.11.296 ALM-44006 Presto Worker Process GC Time Exceeds the Threshold....1850
9.11.297 ALM-45000 HetuEngine Service Unavailable....1852
9.11.298 ALM-45001 Faulty HetuEngine Compute Instances....1856
9.11.299 ALM-45004 Tasks Stacked on HetuEngine Compute Instance....1858
9.11.300 ALM-45005 CPU Usage of HetuEngine Compute Instance Exceeded the Threshold....1861
9.11.301 ALM-45006 Memory Usage of a HetuEngine Compute Instance Exceeded the Threshold....1864

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.302 ALM-45007 Number of Workers of a HetuEngine Compute Instance Is Less Than the Threshold 1866
9.11.303 ALM-45008 Query Latency of HetuEngine Compute Instances Exceeds the Threshold..... 1868
9.11.304 ALM-45009 Task Failure Rate of HetuEngine Compute Instances Exceeds the Threshold..... 1871
9.11.305 ALM-45175 Average Time for Calling OBS Metadata APIs Is Greater than the Threshold..... 1874
9.11.306 ALM-45176 Success Rate of Calling OBS Metadata APIs Is Lower than the Threshold..... 1876
9.11.307 ALM-45177 Success Rate of Calling OBS Data Read APIs Is Lower than the Threshold..... 1877
9.11.308 ALM-45178 Success Rate of Calling OBS Data Write APIs Is Lower Than the Threshold..... 1879
9.11.309 ALM-45179 Number of Failed OBS readFully API Calls Exceeds the Threshold..... 1881
9.11.310 ALM-45180 Number of Failed OBS read API Calls Exceeds the Threshold..... 1883
9.11.311 ALM-45181 Number of Failed OBS write API Calls Exceeds the Threshold..... 1885
9.11.312 ALM-45182 Number of Throttled OBS Operations Exceeds the Threshold..... 1887
9.11.313 ALM-45275 Ranger Service Unavailable..... 1889
9.11.314 ALM-45276 Abnormal RangerAdmin Status..... 1891
9.11.315 ALM-45277 RangerAdmin Heap Memory Usage Exceeds the Threshold..... 1893
9.11.316 ALM-45278 RangerAdmin Direct Memory Usage Exceeds the Threshold..... 1896
9.11.317 ALM-45279 RangerAdmin Non Heap Memory Usage Exceeds the Threshold..... 1898
9.11.318 ALM-45280 RangerAdmin GC Duration Exceeds the Threshold..... 1900
9.11.319 ALM-45281 UserSync Heap Memory Usage Exceeds the Threshold..... 1902
9.11.320 ALM-45282 UserSync Direct Memory Usage Exceeds the Threshold..... 1904
9.11.321 ALM-45283 UserSync Non Heap Memory Usage Exceeds the Threshold..... 1906
9.11.322 ALM-45284 UserSync Garbage Collection (GC) Time Exceeds the Threshold..... 1908
9.11.323 ALM-45285 TagSync Heap Memory Usage Exceeds the Threshold..... 1911
9.11.324 ALM-45286 TagSync Direct Memory Usage Exceeds the Threshold..... 1913
9.11.325 ALM-45287 TagSync Non Heap Memory Usage Exceeds the Threshold..... 1915
9.11.326 ALM-45288 TagSync Garbage Collection (GC) Time Exceeds the Threshold..... 1917
9.11.327 ALM-45289 PolicySync Heap Memory Usage Exceeds the Threshold..... 1919
9.11.328 ALM-45290 PolicySync Direct Memory Usage Exceeds the Threshold..... 1921
9.11.329 ALM-45291 PolicySync Non-Heap Memory Usage Exceeds the Threshold..... 1923
9.11.330 ALM-45292 PolicySync GC Duration Exceeds the Threshold..... 1925
9.11.331 ALM-45293 Ranger User Synchronization Exception..... 1928
9.11.332 ALM-45294 RangerKMS Process Is Abnormal..... 1930
9.11.333 ALM-45425 ClickHouse Service Unavailable..... 1932
9.11.334 ALM-45426 ClickHouse Service Quantity Quota Usage in ZooKeeper Exceeds the Threshold. 1935
9.11.335 ALM-45427 ClickHouse Service Capacity Quota Usage in ZooKeeper Exceeds the Threshold.. 1938
9.11.336 ALM-45428 ClickHouse Disk I/O Exception..... 1940
9.11.337 ALM-45429 Table Metadata Synchronization Failed on the Added ClickHouse Node..... 1943
9.11.338 ALM-45430 Permission Metadata Synchronization Failed on the Added ClickHouse Node..... 1945
9.11.339 ALM-45431 Improper ClickHouse Instance Distribution for Topology Allocation..... 1947
9.11.340 ALM-45432 ClickHouse User Synchronization Process Fails..... 1949
9.11.341 ALM-45433 ClickHouse AZ Topology Exception..... 1951
9.11.342 ALM-45434 A Single Replica Exists in the ClickHouse Data Table..... 1954
9.11.343 ALM-45435 Inconsistent Metadata of ClickHouse Tables..... 1956

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.344 ALM-45436 Skew ClickHouse Table Data....1958
9.11.345 ALM-45437 Excessive Parts in the ClickHouse Table....1960
9.11.346 ALM-45438 ClickHouse Disk Usage Exceeds 80%....1962
9.11.347 ALM-45439 ClickHouse Node Enters the Read-Only Mode....1964
9.11.348 ALM-45440 Inconsistency Between ClickHouse Replicas....1965
9.11.349 ALM-45441 Zookeeper Disconnected....1969
9.11.350 ALM-45442 Too Many Concurrent SQL Statements....1971
9.11.351 ALM-45443 Slow SQL Queries in the Cluster....1973
9.11.352 ALM-45444 Abnormal ClickHouse Process....1975
9.11.353 ALM-45445 Failed to Send Data Files to Remote Shards When ClickHouse Writes Data to a Distributed Table....1977
9.11.354 ALM-45446 Mutation Task of ClickHouse Is Not Complete for a Long Time....1980
9.11.355 ALM-45447 ClickHouse Table Read-Only....1982
9.11.356 ALM-45448 Rapid Increase of Znodes Used by ClickHouse....1985
9.11.357 ALM-45449 The Counter Number of zxid Used by ClickHouse Exceeds the Threshold....1987
9.11.358 ALM-45450 ClickHouse Failed to Obtain a Temporary Agency Credential....1989
9.11.359 ALM-45451 ClickHouse Failed to Access OBS....1990
9.11.360 ALM-45452 ClickHouse's Local Disk Space Is Below the Cold-Hot Separation Threshold....1992
9.11.361 ALM-45475 A Single Replica Exists in the Kudu Data Table....1994
9.11.362 ALM-45476 Number of Tablets of the KuduTServer Process Exceeds the Threshold....1996
9.11.363 ALM-45477 Failed to Restore Data After a Disk of Kudu Is Replaced....1997
9.11.364 ALM-45478 Kudu Failed Data Balancing....2000
9.11.365 ALM-45479 Number of Tablets of the Tserver Process Exceeds the Threshold....2001
9.11.366 ALM-45480 Tablet Leaders of a Tserver Process Are Unevenly Distributed....2003
9.11.367 ALM-45481 KuduTserver Has Full Disks....2006
9.11.368 ALM-45585 IoTDB Service Unavailable....2008
9.11.369 ALM-45586 IoTDBServer Heap Memory Usage Exceeds the Threshold....2010
9.11.370 ALM-45587 IoTDBServer GC Duration Exceeds the Threshold....2012
9.11.371 ALM-45588 IoTDBServer Direct Memory Usage Exceeds the Threshold....2014
9.11.372 ALM-45589 ConfigNode Heap Memory Usage Exceeds the Threshold....2016
9.11.373 ALM-45590 ConfigNode GC Duration Exceeds the Threshold....2018
9.11.374 ALM-45591 ConfigNode Direct Memory Usage Exceeds the Threshold....2020
9.11.375 ALM-45592 IoTDBServer RPC Execution Duration Exceeds the Threshold....2023
9.11.376 ALM-45593 IoTDBServer Flush Execution Duration Exceeds the Threshold....2024
9.11.377 ALM-45594 IoTDBServer Intra-Space Merge Duration Exceeds the Threshold....2026
9.11.378 ALM-45595 IoTDBServer Cross-Space Merge Duration Exceeds the Threshold....2027
9.11.379 ALM-45596 Procedure Execution Failed....2028
9.11.380 ALM-45615 CDL Service Unavailable....2030
9.11.381 ALM-45616 CDL Job Execution Exception....2032
9.11.382 ALM-45617 Data Queued in the CDL Replication Slot Exceeds the Threshold....2035
9.11.383 ALM-45635 FlinkServer Job Execution Failure....2037
9.11.384 ALM-45636 Flink Job Checkpoints Keep Failing....2038
9.11.385 ALM-45636 Number of Consecutive Checkpoint Failures of a Flink Job Exceeds the Threshold2040

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.386 ALM-45637 FlinkServer Task Is Continuously Under Back Pressure....2043
9.11.387 ALM-45638 Number of Restarts After FlinkServer Job Failures Exceeds the Threshold....2045
9.11.388 ALM-45638 Number of Restarts After Flink Job Failures Exceeds the Threshold....2046
9.11.389 ALM-45639 Checkpointing of a Flink Job Times Out....2049
9.11.390 ALM-45640 FlinkServer Heartbeat Interruption Between the Active and Standby Nodes....2051
9.11.391 ALM-45641 Data Synchronization Exception Between the Active and Standby FlinkServer Nodes....2053
9.11.392 ALM-45642 RocksDB Continuously Triggers Write Traffic Limiting....2057
9.11.393 ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold....2061
9.11.394 ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold 2064
9.11.395 ALM-45645 Pending Flush Size of RocksDB Continuously Exceeds the Threshold....2068
9.11.396 ALM-45646 Pending Compaction Size of RocksDB Continuously Exceeds the Threshold....2072
9.11.397 ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold....2075
9.11.398 ALM-45648 RocksDB Frequently Encounters Write-Stopped....2079
9.11.399 ALM-45649 P95 Latency of RocksDB Get Requests Continuously Exceeds the Threshold....2082
9.11.400 ALM-45650 P95 Latency of RocksDB Write Requests Continuously Exceeds the Threshold....2086
9.11.401 ALM-45652 Flink Service Unavailable....2090
9.11.402 ALM-45653 Invalid Flink HA Certificate File....2092
9.11.403 ALM-45654 Flink HA Certificate Is About to Expire....2094
9.11.404 ALM-45655 Flink HA Certificate File Has Expired....2096
9.11.405 ALM-45736 Guardian Service Unavailable....2098
9.11.406 ALM-45737 TokenServer Heap Memory Usage Exceeds the Threshold....2100
9.11.407 ALM-45738 TokenServer Direct Memory Usage Exceeds the Threshold....2102
9.11.408 ALM-45739 TokenServer Non-Heap Memory Usage Exceeds the Threshold....2105
9.11.409 ALM-45740 TokenServer GC Duration Exceeds the Threshold....2107
9.11.410 ALM-45741 Failed to Call the ECS securitykey API....2109
9.11.411 ALM-45742 Failed to Call the ECS Metadata API....2111
9.11.412 ALM-45743 Failed to Call the IAM API....2112
9.11.413 ALM-47001 MemArtsCC Service Unavailable....2113
9.11.414 ALM-47002 MemArtsCC Disk Fault....2115
9.11.415 ALM-47003 Memory Usage of the MemArtsCC Worker Process Exceeds the Threshold....2116
9.11.416 ALM-47004 Average Latency of MemArtsCC Worker Read Requests Exceeds the Threshold....2118
9.11.417 ALM-50201 Doris Service Unavailable....2120
9.11.418 ALM-50202 FE CPU Usage Exceeds the Threshold....2121
9.11.419 ALM-50203 FE Memory Usage Exceeds the Threshold....2123
9.11.420 ALM-50205 BE CPU Usage Exceeds the Threshold....2125
9.11.421 ALM-50206 BE Memory Usage Exceeds the Threshold....2127
9.11.422 ALM-50207 Ratio of Connections to the FE MySQL Port to the Maximum Connections Allowed Exceeds the Threshold....2129
9.11.423 ALM-50208 Failures to Clear Historical Metadata Image Files Exceed the Threshold....2130
9.11.424 ALM-50209 Failures to Generate Doris Metadata Image Files Exceed the Threshold....2132
9.11.425 ALM-50210 Maximum Compaction Score of All BE Nodes Exceeds the Threshold....2134

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.426 ALM-50211 FE Queue Length of BE Periodic Report Tasks Exceeds the Threshold....2136
9.11.427 ALM-50212 Accumulated Old-Generation GC Duration of the FE Process Exceeds the Threshold....2139
9.11.428 ALM-50213 Number of Tasks Queuing in the FE Thread Pool for Interacting with BE Exceeds the Threshold....2140
9.11.429 ALM-50214 Number of Tasks Queuing in the FE Thread Pool for Task Processing Exceeds the Threshold....2142
9.11.430 ALM-50215 Longest Duration of RPC Requests Received by Each FE Thrift Method Exceeds the Threshold....2144
9.11.431 ALM-50216 Memory Usage of the FE Node Exceeds the Threshold....2146
9.11.432 ALM-50217 Heap Memory Usage of the FE Node Exceeds the Threshold....2148
9.11.433 ALM-50219 Length of the Queue in the Thread Pool for Query Execution Exceeds the Threshold....2151
9.11.434 ALM-50220 Error Rate of TCP Packet Receiving Exceeds the Threshold....2153
9.11.435 ALM-50221 BE Data Disk Usage Exceeds the Threshold....2154
9.11.436 ALM-50222 Disk Status of a Specified Data Directory on BE Is Abnormal....2156
9.11.437 ALM-50223 Maximum Memory Required by BE Is Greater Than the Remaining Memory of the Machine....2158
9.11.438 ALM-50224 Failures a Certain Task Type on BE Are Increasing....2160
9.11.439 ALM-50225 FE Instance Fault....2162
9.11.440 ALM-50226 BE Instance Fault....2163
9.11.441 ALM-50401 Number of JobServer Jobs Waiting to Be Executed Exceeds the Threshold....2165
9.11.442 ALM-50402 JobGateway Service Unavailable....2167
9.11.443 ALM-50406 Failure Rate of the JobServer Job Submission API Exceeds the Threshold....2169
9.11.444 ALM-50407 Failure Rate of the JobServer Job Query API Exceeds the Threshold....2170
9.11.445 ALM-50408 Failure Rate of the JobServer Job Termination API Exceeds the Threshold....2172
9.11.446 ALM-12001 Audit Log Dump Failure (For MRS 2.*x* or Earlier)....2174
9.11.447 ALM-12002 HA Resource Abnormal (For MRS 2.*x* or Earlier)....2175
9.11.448 ALM-12004 OLdap Resource Abnormal (For MRS 2.*x* or Earlier)....2178
9.11.449 ALM-12005 OKerberos Resource Abnormal (For MRS 2.*x* or Earlier)....2179
9.11.450 ALM-12006 Node Fault (For MRS 2.*x* or Earlier)....2181
9.11.451 ALM-12007 Process Fault (For MRS 2.*x* or Earlier)....2182
9.11.452 ALM-12010 Manager Heartbeat Interruption Between the Active and Standby Nodes (For MRS 2.*x* or Earlier)....2184
9.11.453 ALM-12011 Data Synchronization Exception Between the Active and Standby Manager Nodes (For MRS 2.*x* or Earlier)....2186
9.11.454 ALM-12012 NTP Service Abnormal (For MRS 2.*x* or Earlier)....2187
9.11.455 ALM-12014 Device Partition Lost (For MRS 2.*x* or Earlier)....2190
9.11.456 ALM-12015 Device Partition File System Read-Only (For MRS 2.*x* or Earlier)....2192
9.11.457 ALM-12016 CPU Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2194
9.11.458 ALM-12017 Insufficient Disk Capacity (For MRS 2.*x* or Earlier)....2195
9.11.459 ALM-12018 Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2198
9.11.460 ALM-12027 Host PID Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2199
9.11.461 ALM-12028 Number of Processes in the D State on the Host Exceeds the Threshold (For MRS 2.*x* or Earlier)....2201

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.462 ALM-12031 User omm or Password Is About to Expire (For MRS 2.*x* or Earlier)....2203
9.11.463 ALM-12032 User ommdba or Password Is About to Expire (For MRS 2.*x* or Earlier)....2204
9.11.464 ALM-12033 Slow Disk Fault (For MRS 2.*x* or Earlier)....2206
9.11.465 ALM-12034 Periodic Backup Failure (For MRS 2.*x* or Earlier)....2213
9.11.466 ALM-12035 Unknown Data Status After Recovery Task Failure (For MRS 2.*x* or Earlier)....2215
9.11.467 ALM-12037 NTP Server Abnormal (For MRS 2.*x* or Earlier)....2216
9.11.468 ALM-12038 Monitoring Indicator Dump Failure (For MRS 2.*x* or Earlier)....2218
9.11.469 ALM-12039 GaussDB Data Is Not Synchronized (For MRS 2.*x* or Earlier)....2220
9.11.470 ALM-12040 Insufficient System Entropy (For MRS 2.*x* or Earlier)....2223
9.11.471 ALM-12041 Permission of Key Files Is Abnormal (For MRS 2.*x* or Earlier)....2224
9.11.472 ALM-12042 Key File Configurations Are Abnormal (For MRS 2.*x* or Earlier)....2226
9.11.473 ALM-12043 DNS Parsing Duration Exceeds the Threshold (For MRS 2.*x* or Earlier)....2227
9.11.474 ALM-12045 Read Packet Dropped Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)....2230
9.11.475 ALM-12046 Write Packet Dropped Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)....2234
9.11.476 ALM-12047 Read Packet Error Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)....2236
9.11.477 ALM-12048 Write Packet Error Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)....2237
9.11.478 ALM-12049 Read Throughput Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)....2239
9.11.479 ALM-12050 Write Throughput Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)....2241
9.11.480 ALM-12051 Disk Inode Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2243
9.11.481 ALM-12052 Usage of Temporary TCP Ports Exceeds the Threshold (For MRS 2.*x* or Earlier)....2245
9.11.482 ALM-12053 File Handle Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2247
9.11.483 ALM-12054 Invalid Certificate File (For MRS 2.*x* or Earlier)....2249
9.11.484 ALM-12055 Certificate File Is About to Expire (For MRS 2.*x* or Earlier)....2251
9.11.485 ALM-12180 Disk Card I/O (For MRS 2.*x* or Earlier)....2254
9.11.486 ALM-12357 Failed to Export Audit Logs to OBS (For MRS 2.*x* or Earlier)....2257
9.11.487 ALM-13000 ZooKeeper Service Unavailable (For MRS 2.*x* or Earlier)....2259
9.11.488 ALM-13001 Available ZooKeeper Connections Are Insufficient (For MRS 2.*x* or Earlier)....2261
9.11.489 ALM-13002 ZooKeeper Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2264
9.11.490 ALM-14000 HDFS Service Unavailable (For MRS 2.*x* or Earlier)....2265
9.11.491 ALM-14001 HDFS Disk Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2267
9.11.492 ALM-14002 DataNode Disk Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)....2269
9.11.493 ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold (For MRS 2.*x* or Earlier)....2271
9.11.494 ALM-14004 Number of Damaged HDFS Blocks Exceeds the Threshold (For MRS 2.*x* or Earlier)....2272
9.11.495 ALM-14006 Number of HDFS Files Exceeds the Threshold (For MRS 2.*x* or Earlier)....2274
9.11.496 ALM-14007 HDFS NameNode Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)2275
9.11.497 ALM-14008 HDFS DataNode Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)2277
9.11.498 ALM-14009 Number of Faulty DataNodes Exceeds the Threshold (For MRS 2.*x* or Earlier)....2278
9.11.499 ALM-14010 NameService Is Abnormal (For MRS 2.*x* or Earlier)....2280
9.11.500 ALM-14011 HDFS DataNode Data Directory Is Not Configured Properly (For MRS 2.*x* or Earlier)....2283
9.11.501 ALM-14012 HDFS Journalnode Data Is Not Synchronized (For MRS 2.*x* or Earlier)....2286

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.502 ALM-16000 Percentage of Sessions Connected to the HiveServer to the Maximum Number Allowed Exceeds the Threshold (For MRS 2.*x* or Earlier) 2288
9.11.503 ALM-16001 Hive Warehouse Space Usage Exceeds the Threshold (For MRS 2.*x* or Earlier) 2290
9.11.504 ALM-16002 Hive SQL Execution Success Rate Is Lower Than the Threshold (For MRS 2.*x* or Earlier) 2292
9.11.505 ALM-16004 Hive Service Unavailable (For MRS 2.*x* or Earlier) 2294
9.11.506 ALM-16005 Number of Failed Hive SQL Executions in the Last Period Exceeds the Threshold (For MRS 2.*x* or Earlier) 2298
9.11.507 ALM-18000 Yarn Service Unavailable (For MRS 2.*x* or Earlier) 2299
9.11.508 ALM-18002 NodeManager Heartbeat Lost (For MRS 2.*x* or Earlier) 2301
9.11.509 ALM-18003 NodeManager Unhealthy (For MRS 2.*x* or Earlier) 2302
9.11.510 ALM-18004 NodeManager Disk Usability Ratio Is Lower Than the Threshold (For MRS 2.*x* or Earlier) 2303
9.11.511 ALM-18006 MapReduce Job Execution Timeout (For MRS 2.*x* or Earlier) 2304
9.11.512 ALM-18008 Heap Memory Usage of Yarn ResourceManager Exceeds the Threshold (For MRS 2.*x* or Earlier) 2306
9.11.513 ALM-18009 Heap Memory Usage of MapReduce JobHistoryServer Exceeds the Threshold (For MRS 2.*x* or Earlier) 2308
9.11.514 ALM-18010 Number of Pending Yarn Tasks Exceeds the Threshold (For MRS 2.*x* or Earlier) 2310
9.11.515 ALM-18011 Memory of Pending Yarn Tasks Exceeds the Threshold (For MRS 2.*x* or Earlier) 2311
9.11.516 ALM-18012 Number of Terminated Yarn Tasks in the Last Period Exceeds the Threshold (For MRS 2.*x* or Earlier) 2313
9.11.517 ALM-18013 Number of Failed Yarn Tasks in the Last Period Exceeds the Threshold (For MRS 2.*x* or Earlier) 2314
9.11.518 ALM-19000 HBase Service Unavailable (For MRS 2.*x* or Earlier) 2315
9.11.519 ALM-19006 HBase Replication Sync Failed (For MRS 2.*x* or Earlier) 2316
9.11.520 ALM-19007 HBase Merge Queue Exceeds the Threshold (for 2.x and Earlier Versions) 2319
9.11.521 ALM-20002 Hue Service Unavailable (For MRS 2.*x* or Earlier) 2321
9.11.522 ALM-23001 Loader Service Unavailable (For MRS 2.*x* or Earlier) 2323
9.11.523 ALM-24000 Flume Service Unavailable (For MRS 2.*x* or Earlier) 2327
9.11.524 ALM-24001 Flume Agent Is Abnormal (For MRS 2.*x* or Earlier) 2328
9.11.525 ALM-24003 Flume Client Connection Interrupted (For MRS 2.*x* or Earlier) 2330
9.11.526 ALM-24004 Flume Fails to Read Data (For MRS 2.*x* or Earlier) 2332
9.11.527 ALM-24005 Data Transmission by Flume Is Abnormal (For MRS 2.*x* or Earlier) 2334
9.11.528 ALM-25000 LdapServer Service Unavailable (For MRS 2.*x* or Earlier) 2336
9.11.529 ALM-25004 Abnormal LdapServer Data Synchronization (For MRS 2.*x* or Earlier) 2338
9.11.530 ALM-25500 KrbServer Service Unavailable (For MRS 2.*x* or Earlier) 2341
9.11.531 ALM-26051 Storm Service Unavailable (For MRS 2.*x* or Earlier) 2343
9.11.532 ALM-26052 Number of Available Supervisors in Storm Is Lower Than the Threshold (For MRS 2.*x* or Earlier) 2345
9.11.533 ALM-26053 Slot Usage of Storm Exceeds the Threshold (For MRS 2.*x* or Earlier) 2346
9.11.534 ALM-26054 Heap Memory Usage of Storm Nimbus Exceeds the Threshold (For MRS 2.*x* or Earlier) 2348
9.11.535 ALM-27001 DBService Unavailable (For MRS 2.*x* or Earlier) 2350

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.536 ALM-27003 DBService Heartbeat Interruption Between the Active and Standby Nodes (For MRS 2.*x* or Earlier) 2352
9.11.537 ALM-27004 Data Inconsistency Between Active and Standby DBServices (For MRS 2.*x* or Earlier) 2354
9.11.538 ALM-28001 Spark Service Unavailable (For MRS 2.*x* or Earlier) 2356
9.11.539 ALM-38000 Kafka Service Unavailable (For MRS 2.*x* or Earlier) 2358
9.11.540 ALM-38001 Insufficient Kafka Disk Capacity (For MRS 2.*x* or Earlier) 2360
9.11.541 ALM-38002 Heap Memory Usage of Kafka Exceeds the Threshold (For MRS 2.*x* or Earlier) 2363
9.11.542 ALM-43001 Spark Service Unavailable (For MRS 2.*x* or Earlier) 2364
9.11.543 ALM-43006 Heap Memory Usage of the JobHistory Process Exceeds the Threshold (For MRS 2.*x* or Earlier) 2366
9.11.544 ALM-43007 Non-Heap Memory Usage of the JobHistory Process Exceeds the Threshold (For MRS 2.*x* or Earlier) 2367
9.11.545 ALM-43008 Direct Memory Usage of the JobHistory Process Exceeds the Threshold (For MRS 2.*x* or Earlier) 2369
9.11.546 ALM-43009 JobHistory GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier) 2370
9.11.547 ALM-43010 Heap Memory Usage of the JDBCServer Process Exceeds the Threshold (For MRS 2.*x* or Earlier) 2372
9.11.548 ALM-43011 Non-Heap Memory Usage of the JDBCServer Process Exceeds the Threshold (For MRS 2.*x* or Earlier) 2373
9.11.549 ALM-43012 Direct Memory Usage of the JDBCServer Process Exceeds the Threshold (For MRS 2.*x* or Earlier) 2375
9.11.550 ALM-43013 JDBCServer GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier) 2376
9.11.551 ALM-44004 Presto Coordinator Resource Group Queuing Tasks Exceed the Threshold (For MRS 2.*x* or Earlier) 2378
9.11.552 ALM-44005 Presto Coordinator Process GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier) 2379
9.11.553 ALM-44006 Presto Worker Process GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier) 2380
9.11.554 ALM-45325 Presto Service Unavailable (For MRS 2.*x* or Earlier) 2382
9.12 Configuring Remote O&M for an MRS Cluster 2383
9.13 Common Ports for MRS Cluster Services 2384

10 Configuring Storage-Compute Decoupling for an MRS Cluster 2402

10.1 MRS Storage-Compute Decoupling Overview 2402
10.2 Interconnecting an MRS Cluster with OBS Through Guardian 2405
10.2.1 Disabling Ranger OBS Path Authentication for Guardian 2406
10.2.2 Enabling Ranger OBS Path Authentication for Guardian 2411
10.2.3 Example for Interconnecting an MRS Cluster with OBS 2417
10.2.3.1 Accessing OBS Using Flink Through Guardian 2417
10.2.3.2 Accessing OBS Using HDFS Through Guardian 2419
10.2.3.3 Accessing OBS Using HetuEngine Through Guardian 2421
10.2.3.4 Accessing OBS Using Hive Through Guardian 2423
10.2.3.5 Accessing OBS Using Hudi Through Guardian 2426
10.2.3.6 Accessing OBS Using MapReduce Through Guardian 2428
10.2.3.7 Accessing OBS Using Spark Through Guardian 2429
10.2.3.8 Accessing OBS Using YARN Through Guardian 2434

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

10.3 Interconnecting an MRS Cluster with OBS Using an IAM Agency 2436
10.3.1 Interconnecting an MRS Cluster with OBS Using an IAM Agency 2436
10.3.2 Configuring the Policy for Clearing Recycle Bin Directories of MRS Cluster Components 2444
10.3.3 Example for Interconnecting an MRS Cluster with OBS 2448
10.3.3.1 Interconnecting Flink with OBS Using an IAM Agency 2448
10.3.3.2 Interconnecting Flume with OBS Using an IAM Agency 2450
10.3.3.3 Interconnecting HDFS with OBS Using an IAM Agency 2451
10.3.3.4 Interconnecting Hive with OBS Using an IAM Agency 2452
10.3.3.5 Interconnecting Hudi with OBS Using an IAM Agency 2456
10.3.3.6 Interconnecting MapReduce with OBS Using an IAM Agency 2457
10.3.3.7 Interconnecting Presto with OBS Using an IAM Agency 2459
10.3.3.8 Interconnecting Spark with OBS Using an IAM Agency 2461
10.3.3.9 Interconnecting Sqoop with OBS Using an IAM Agency 2465
10.3.4 Configuring Fine- Grained OBS Access Permissions for MRS Cluster Users 2466
10.3.5 How Do I Access OBS Using an MRS Client Installed Outside a Cluster? 2472

11 FAQs 2475

11.1 MRS Basics 2475
11.1.1 What Is MRS Used For? 2475
11.1.2 What Types of Distributed Storage Does MRS Support? 2475
11.1.3 What Are Regions and AZs? 2476
11.1.4 Can I Change the Network Segment of Nodes in an MRS Cluster? 2476
11.1.5 Can I Downgrade the Specifications of an MRS Cluster Node? 2477
11.1.6 Are Hive Components of Different Versions Compatible with Each Other? 2477
11.1.7 What Are the Differences Between OBS and HDFS in Data Storage? 2477
11.1.8 What Are the Solutions for Processing over 1 Billion Data Records? 2478
11.1.9 What Are the Advantages of ZSTD? 2478
11.2 Cluster Creation 2478
11.2.1 How Do I Create an MRS Cluster Using a Custom Security Group? 2478
11.2.2 What Should I Do If HDFS, Yarn, and MapReduce Components Are Unavailable When I Create an MRS Cluster? 2478
11.2.3 What Should I Do If the ZooKeeper Component Is Unavailable When I Create an MRS Cluster? 2479
11.3 Web Page Access 2479
11.3.1 How Do I Change the Session Timeout Duration for an Open Source Component Web UI? 2479
11.3.2 What Can I Do If the Dynamic Resource Plan Page in MRS Tenant Management Cannot Be Refreshed? 2481
11.3.3 What Do I Do If the Kafka Topic Monitoring Tab Is Not Displayed on Manager? 2482
11.3.4 Why Can't I Find the User Management Page on MRS Manager? 2482
11.3.5 What Should I Do If the Excel File Downloaded by Hue Cannot Be Opened? 2482
11.4 Authentication and Permission 2484
11.4.1 What Is the User for Logging in to FusionInsight Manager? 2484
11.4.2 How Do I Query and Change the Password Validity Period of a User In a Cluster? 2484
11.4.3 Does an MRS Cluster Support Access Permission Control If Kerberos Authentication Is not Enabled? 2485

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.4.4 How Do I Add Tenant Management Permission to Users in a Cluster?.....2486
11.4.5 Can I Configure Account Permissions on the Hue Web UI?.....2486
11.4.6 How Do I View the Hive Table Created by Another User?.....2487
11.4.7 How Do I Prevent Kerberos Authentication Expiration?.....2488
11.4.8 How Do I Enable or Disable Kerberos Authentication for an Existing MRS Cluster?.....2489
11.4.9 What Are the Ports of the Kerberos Authentication Service?.....2489
11.5 Client Usage.....2490
11.5.1 How Do I Disable SASL Authentication for ZooKeeper?.....2490
11.5.2 What Can I Do If the Error Message "Permission denied" Is Displayed When kinit Is Executed on a Client Outside the MRS Cluster?.....2490
11.5.3 What Should I Do If an Alarm Is Reported Indicating that the Memory Is Insufficient When I Execute a SQL Statement on the ClickHouse Client?.....2491
11.5.4 How Do I Connect to Spark Shell from MRS?.....2491
11.5.5 How Do I Connect to Spark Beeline from MRS?.....2491
11.5.6 What Should I Do If the Connection to the ClickHouse Server Fails and Error Code 516 Is Reported?.....2492
11.6 Component Configurations.....2493
11.6.1 Does MRS Support Running Hive on Kudu?.....2493
11.6.2 Does an MRS Cluster Support Hive on Spark?.....2493
11.6.3 Can I Change the IP address of DBService?.....2493
11.6.4 What Access Protocols Does Kafka Support?.....2493
11.6.5 What Python Versions Are Supported by Spark Tasks in an MRS Cluster?.....2494
11.6.6 What Are the Restrictions on the Storm Log Size in an MRS 2.1.0 Cluster?.....2494
11.6.7 How Do I Modify the HDFS fs.defaultFS of an Existing Cluster?.....2494
11.6.8 Can MRS Run Multiple Flume Tasks at a Time?.....2494
11.6.9 How Do I Change FlumeClient Logs to Standard Logs?.....2495
11.6.10 Where Are the JAR Files and Environment Variables of Hadoop Stored?.....2495
11.6.11 How Do I View HBase Logs?.....2495
11.6.12 How Do I Set the TTL for an HBase Table?.....2495
11.6.13 How Do I Change the Number of HDFS Replicas?.....2495
11.6.14 How Do I Modify the HDFS Active/Standby Switchover Class?.....2496
11.6.15 What Data Type in Hive Tables Is Recommended for the Number Type of DynamoDB?.....2496
11.6.16 Can I Export the Query Result of Hive Data?.....2496
11.6.17 What Should I Do If an Error Occurs When Hive Runs the **beeline -e** Command to Execute Multiple Statements?.....2497
11.6.18 What Do I Do If "over max user connections" Is Displayed When Hue Connects to HiveServer?2497
11.6.19 How Do I View MRS Hive Metadata?.....2498
11.6.20 How Do I Reset Kafka Data?.....2498
11.6.21 What Should I Do If the Error Message "Not Authorized to access group XXX" Is Displayed When Kafka Topics Are Consumed?.....2499
11.6.22 What Compression Algorithms Does Kudu Support?.....2499
11.6.23 How Do I View Kudu Logs?.....2499
11.6.24 How Do I Handle the Kudu Service Exceptions Generated During Cluster Creation?.....2499
11.6.25 How Do I Configure Other Data Sources on Presto?.....2500

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.6.26 How Do I Specify a Log Path When Submitting a Task in an MRS Storm Cluster?.....2502
11.6.27 How Do I Check the ResourceManager Configuration of Yarn?.....2502
11.6.28 How Do I Modify the allow_drop_detached Parameter of ClickHouse?.....2503
11.6.29 How Do I Add a Policy for Periodically Deleting Large ClickHouse System Table Logs?.....2504
11.7 Cluster Management.....2505
11.7.1 How Do I View All Clusters?.....2505
11.7.2 How Do I View MRS Operation Logs?.....2505
11.7.3 How Do I View MRS Cluster Configuration Information?.....2506
11.7.4 How Do I Add Components to an MRS Cluster?.....2506
11.7.5 How Do I Cancel Message Notifications for MRS Cluster Alarms?.....2506
11.7.6 Why Is the Resource Pool Memory Displayed in the MRS Cluster Smaller Than the Actual Cluster Memory?.....2506
11.7.7 What Is the Python Version Installed for an MRS Cluster?.....2507
11.7.8 What Can I Do If the Time Information of an MRS Cluster Node Is Incorrect?.....2507
11.7.9 What Are the Differences and Relationships Between the MRS Management Console and MRS Manager?.....2507
11.7.10 How Do I Disable the Firewall Service on Nodes in an MRS Cluster?.....2509
11.7.11 How Do I Switch the Login Mode of a Node in an MRS Cluster?.....2509
11.7.12 How Do I Access an MRS Cluster from a Node Outside the Cluster?.....2510
11.7.13 Does Kafka Topic Monitoring in MRS Clusters Support Alarm Reporting?.....2511
11.7.14 How Do I View Running Resource Queues When an Alarm Indicating Insufficient YARN Queue Resources Is Generated?.....2511
11.7.15 How Do I Understand the Multi-Level Chart Statistics in the HBase Operation Requests Metric?.....2511
11.8 Node Management.....2512
11.8.1 Can I Change MRS Cluster Nodes on the MRS Console?.....2512
11.8.2 How Do I Query the Startup Time of an MRS Node?.....2512
11.8.3 What Do I Do If Trust Relationships Between Nodes Are Abnormal?.....2512
11.8.4 Can Sudo Logs of Nodes in an MRS Cluster Be Cleared?.....2513
11.8.5 Does an MRS Cluster Support System Reinstallation?.....2514
11.8.6 Can I Change the OS of an MRS Cluster?.....2514
11.9 Component Management.....2514
11.9.1 Can I Delete Components Installed in an MRS Cluster?.....2514
11.9.2 How Do I View the Configuration File Directory of Each Component?.....2515
11.9.3 Will Upper-Layer Services Be Affected If the Hive Service Status Is Partially Healthy?.....2515
11.9.4 How Do I Obtain the IP Address and Port Number of a ZooKeeper Instance in the MRS Cluster?2516
11.10 Job Management.....2516
11.10.1 What Types of Spark Jobs Can Be Submitted in a Cluster?.....2517
11.10.2 How Do I Enable Different Service Programs to Use Different Yarn Queues?.....2517
11.10.3 What Should I Do If a Job Fails to Be Submitted and the Error Is Related to OBS?.....2520
11.10.4 Can I Run Multiple Spark Tasks at the Same Time After the Minimum Tenant Resources of an MRS Cluster Is Changed to 0?.....2521
11.10.5 What Are the Differences Between the Client Mode and Cluster Mode of Spark Jobs?.....2521
11.10.6 How Do I View MRS Job Logs?.....2521

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.7 What Can I Do If the System Displays a Message Indicating that the Current User Does Not Exist on Manager When I Submit a Job? 2522
11.10.8 What Can I Do If LauncherJob Fails to Be Executed and the Error Message "jobPropertiesMap is null" Is Displayed? 2523
11.10.9 What Should I Do If the Flink Job Status on the MRS Console Is Inconsistent with That on YARN? 2523
11.10.10 What Should I Do If Error Code 403 Is Reported When a Spark Streaming Job Is Running? 2523
11.10.11 What Should I Do If Error Message "java.io.IOException: Connection reset by peer" Is Displayed During the Execution of a Spark Job? 2524
11.10.12 What Should I Do If the Error Message "requestId=XXX" Is Displayed When a Spark Job Accesses OBS? 2524
11.10.13 What Should I Do If the Error Message "UnknownScannerExeception" Is Displayed for Spark Jobs? 2524
11.10.14 What Can I Do If DataArts Studio Occasionally Fails to Schedule Spark Jobs? 2525
11.10.15 What Should I Do If a Flink Job Fails to Execute and the Error Message "java.lang.NoSuchFieldError: SECURITY_SSL_ENCRYPT_ENABLED" Is Displayed? 2525
11.10.16 What Should I Do If Submitted Yarn Jobs Cannot Be Viewed on the Web UI? 2525
11.10.17 What Can I Do If launcher-job Is Terminated by Yarn When a Flink Task Is Submitted? 2526
11.10.18 What Should I Do If the Error Message "slot request timeout" Is Displayed When I Submit a Flink Job? 2526
11.10.19 FAQs About Importing and Exporting Data Using DistCP Jobs 2526
11.10.20 How Do I View SQL Statements of Hive Jobs on the Yarn Web UI? 2527
11.10.21 How Do I View Logs of a Specified Yarn Task? 2528
11.10.22 What Should I Do If a HiveSQL/HiveScript Job Fails to be Submitted After Hive Is Added? 2528
11.10.23 Where Are the Execution Logs of Spark Jobs Stored? 2528
11.10.24 What Should I Do If an Alarm Indicating Insufficient Memory Is Reported During Spark Task Execution? 2529
11.10.25 What Should I Do If It Takes a Long Time for Spark SQL to Access Hive Partitioned Tables Before a Job Starts? 2529
11.10.26 How Does the System Select the Queue When an MRS Cluster User Is Associated to Multiple Queues? 2530
11.11 Performance Tuning 2530
11.11.1 How Do I Obtain the Hadoop Pressure Test Tool? 2530
11.11.2 How Do I Improve the Resource Utilization of Core Nodes in a Cluster? 2531
11.11.3 How Do I Configure the Knox Memory for an MRS Cluster? 2531
11.11.4 How Do I Adjust the Memory Size of the manager-executor Process? 2532
11.11.5 What Should I Do If the spark.yarn.executor.memoryOverhead Setting Does Not Take Effect? 2532
11.12 Application Development 2533
11.12.1 How Do I Get My Data into OBS or HDFS? 2533
11.12.2 Can MRS Write Data to HBase Through an HBase External Table of Hive? 2534
11.12.3 Does OpenTSDB Support Python APIs? 2534
11.12.4 How Do I Configure the node_id Parameter When Calling the API for Adjusting Cluster Nodes? 2534
11.12.5 How Do I Manage and Use Third-Party JAR Packages for MRS Cluster Components? 2534
11.13 Upgrade and Patching 2535

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.13.1 How Do I Upgrade an MRS Cluster?.....2535
11.13.2 Does MRS Support Kernel Version Upgrade of Components in a Cluster?.....2535

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1 Product Introduction

1.1 What Is MRS?

Big data presents both exciting opportunities and a huge challenge. As the data volume and types increase rapidly, conventional data processing technologies, such as standalone storage systems and relational databases, are struggling to keep up. Rising to this challenge, the Apache Software Foundation (ASF) launched an open source project called Hadoop. Hadoop is an open source distributed computing platform that can fully utilize the computing and storage capabilities of large compute clusters to process massive amounts of data. Hadoop is a powerful framework, but it is not easy to deploy and operationalize — If enterprises try to deploy Hadoop systems all by themselves, they may encounter problems such as high costs, long rollout, difficult maintenance, and inflexible use.

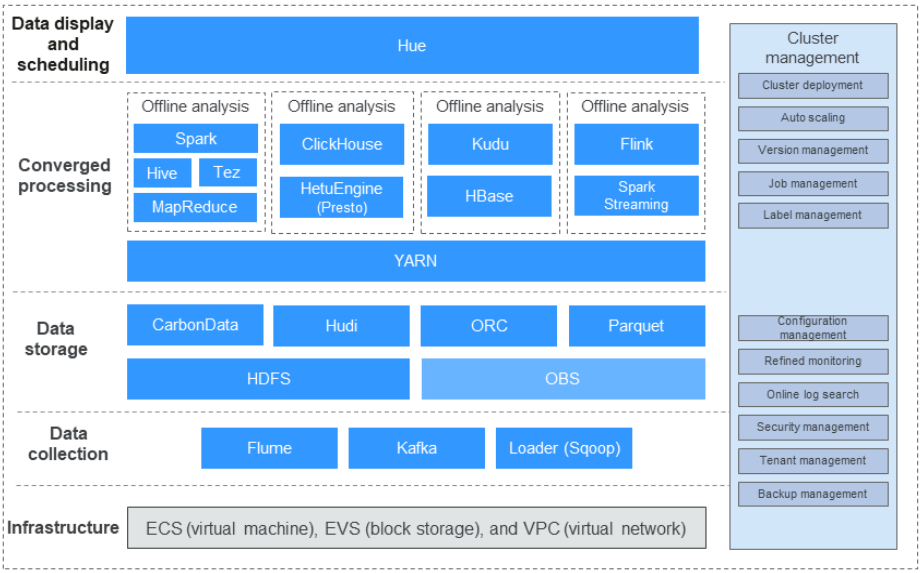
The MapReduce Service (MRS) offers a one-stop service that helps you quickly deploy and manage Hadoop systems on the cloud with ease. With MRS, you can create an enterprise-class Hadoop cluster with just a few clicks of your mouse. Tenants have total control over their Hadoop clusters and can effortlessly run big data components such as Storm, Hadoop, Spark, HBase, and Kafka. MRS supports a full range of open source APIs. By leveraging the cloud platform's deep expertise in compute, storage, and big data, it offers customers a full-stack big data platform featuring high performance, high cost-effectiveness, flexibility, and ease-of-use. Furthermore, the platform can be easily customized to meet new requirements and help enterprises quickly build a massive data processing system and discover new value and business opportunities by analyzing and mining massive amounts of data in real time or in non-real time.

Product Architecture

Figure 1-1 shows the MRS logical architecture.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-1 MRS architecture



MRS includes the infrastructure and an end-to-end big data processing pipeline.

- Infrastructure

  MRS big data clusters fully utilize the high scalability, reliability, and security features of the virtualization layer powered by the cloud platform.

  - Virtual Private Cloud (VPC) provides virtual private networks for each tenant on the cloud.
  - Elastic Volume Service (EVS) provides reliable and high-performance storage.
  - Elastic Cloud Server (ECS) provides VMs that are easily scalable. It works with VPCs, security groups, and the EVS multi-replica mechanism to build an efficient, reliable, and secure computing environment.

- Data collection

  The data collection layer provides the ability to efficiently ingest data from various data sources. It consists of Flume (data ingestion), Loader (relational data loading), and Kafka (highly reliable message queue). Alternatively, you can use Cloud Data Migration (CDM) service to ingest external data to MRS clusters.

- Data storage

  MRS clusters can store both structured and unstructured data. They support multiple efficient data formats to meet the requirements of different computing engines, including:

  - HDFS, which is a general-purpose distributed file system for big data platforms.
  - OBS is an object storage service that features high availability and low cost.

- Converged data processing

  - MRS supports multiple mainstream compute engines, including MapReduce (batch processing), Tez (DAG model), Spark (in-memory computing), Spark Streaming (micro-batch stream computing), Storm

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

(stream computing), and Flink (stream computing). They convert data structures and logic into data models that meet the needs of a variety of big data applications.

- Based on preset data models and easy-to-use SQL data analysis, users can choose Hive (data warehouse), SparkSQL, and Presto (interactive query engine) to run different types of analytical tasks.

- Data display and scheduling

  Data analysis results are displayed intuitively. MRS also integrates with DataArts Studio to provide a one-stop, collaborative big data development platform, helping you easily run a range of different tasks, such as data modeling, data integration, script development, job scheduling, and O&M monitoring, making big data more accessible than ever before.

- Cluster management

  All components of the Hadoop-based big data ecosystem are deployed in distributed mode, and their deployment, management, and O&M are complex.

  MRS provides a unified O&M and management platform for cluster management, supporting one-click cluster deployment, multi-version selection, as well as manual scaling and auto scaling of clusters with zero service interruption. In addition, MRS provides job management, resource tag management, and O&M covering all of the Hadoop components. One-stop O&M capabilities include monitoring, alarm reporting, parameter configuration, and patch upgrade.

Product Advantages

MRS has a strong Hadoop kernel team and is built on top of enterprise-class FusionInsight big data platform. MRS can guarantee multi-level Service Level Agreements (SLAs).

MRS has the following advantages:

- High performance

  MRS supports CarbonData storage solution. CarbonData allows a single copy of data to be used for multiple tasks. It supports features such as multi-level indexing, dictionary encoding, pre-aggregation, dynamic partitioning, and quasi-real-time data query. These features improve I/O scanning and computing performance, allowing tens of billions of data records to be analyzed in seconds. In addition, MRS supports the Superior Scheduler, which outperforms open-source schedulers in every way and enables efficient scheduling in super large clusters (up to 10,000 nodes).

- Cost-effectiveness

  MRS supports a heterogeneous compute and storage infrastructure with decoupled storage and compute, offering a cost-effective mass storage solution. MRS supports fast auto scaling to accommodate changing demand, maximizing resource utilization for customers. MRS clusters can be quickly created and scaled out as you needed, and can be deleted or scaled when you no longer need them.

- High security

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS provides enterprise-class multi-tenant permissions management and security management, with support for table-based and column-based access control and data encryption.

- Easy O&M

  MRS provides an efficient big data cluster management platform that supports one-click rolling patch updates, which ensure the continuity of your services.

- High reliability

  Tested and proven in numerous projects, the long-term reliability and stability of MRS in large-scale deployments can meet enterprise-class standards for production systems. In addition, MRS supports automatic data backup across AZs and regions, as well as automatic anti-affinity, allowing mission-critical VMs to be distributed on different physical machines.

1.2 Advantages

MRS provides enterprise-level big data clusters on the cloud. Tenants can fully control the clusters and run big data components such as Hadoop, Spark, HBase, Kafka, and ClickHouse with ease. MRS frees you from hardware purchase and maintenance. MRS is built based on the enterprise-class big data platform and has been deployed on tens of thousands of nodes in the industry, providing multi-level SLA assurance with professional Hadoop kernel service support. Compared with self-built Hadoop clusters, MRS has the following advantages:

1. **MRS supports one-click cluster creation, deletion, and scaling. You can use an elastic IP address (EIP) to access MRS Manager, making big data clusters easier to use.**
   - Self-built big data clusters pose problems such as high costs, long periods, difficult and inflexible O&M. To solve these problems, MRS provides one-click cluster creation, deletion, scale-out, and scale-in, allowing you to customize the cluster type, component range, number of nodes of each type, VM specifications, availability zones (AZs), VPC network, and authentication information. MRS can automatically create a cluster that meets the configuration requirements. In addition, you can quickly create multi-application clusters, for example, Hadoop analysis cluster, HBase cluster, and Kafka cluster. MRS supports heterogeneous cluster deployment. That is, VMs of different specifications can be combined in a cluster based on CPU types, disk capacities, disk types, and memory sizes.
   - MRS provides an EIP-based secure channel for you to easily access the web UIs of components. This is more convenient than binding an EIP by yourself, and you can access the web UIs with a few clicks, avoiding the steps for logging in to a VPC, adding security group rules, and obtaining a public IP address.
   - MRS provides custom bootstrap actions to flexibly configure your dedicated clusters. Third-party software that is not supported by MRS can be automatically installed, allowing you to perform custom operations such as modifying the cluster running environment.
   - MRS, with the WrapperFS feature and the OBS translation capability (that is, access to OBS through address mapping), can smoothly migrate data

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

from HDFS to OBS. After migration, you can access the data stored in OBS from clients without modifying service code logic.

2. **MRS supports auto scaling, which is more cost-effective than the self-built Hadoop cluster.**

   MRS supports auto scaling to address peak and off-peak service loads. It applies for extra resources during peak hours and releases idle resources during off-peak hours, helping you save idle resources on the big data platform during off-peak hours, minimize costs, and focus on core services.

   In big data applications, especially in periodic data analysis and processing, cluster computing resources need to be dynamically adjusted based on service data changes to meet service requirements. The auto scaling function of MRS enables clusters to be elastically scaled out or in based on cluster loads. In addition, if the data volume changes regularly and you want to scale out or in a cluster before the data volume changes, you can use the MRS resource plan feature. MRS supports two types of auto scaling policies: auto scaling rules and resource plans

   - Auto scaling rules: You can increase or decrease Task nodes based on real-time cluster loads. Auto scaling will be triggered when the data volume changes but there may be some delay.
   - Resource plans: If the data volume changes periodically, you can create resource plans to resize the cluster before the data volume changes, thereby avoiding a delay in increasing or decreasing resources.

   Both auto scaling rules and resource plans can trigger auto scaling. You can configure both of them or configure one of them. Configuring both resource plans and auto scaling rules improves the cluster node scalability to cope with occasionally unexpected data volume peaks.

3. **MRS supports storage-compute decoupling, greatly improving the resource utilization of big data clusters.**

   In the traditional big data architecture where storage and compute resources are integrated, scaling-out is difficult and resources are not well-utilized. To solve these problems, MRS adopts a compute-storage separation architecture. Based on OBS, the storage achieves 99.999999999% reliability and unlimited capacity, supporting continuous growth of enterprise data. Computing resources can be elastically scaled in or out from 0 to *N* nodes. Hundreds of nodes can be quickly provisioned. With the new architecture, compute nodes can be elastically scaled. OBS-based cross-AZ data storage ensures higher reliability, frees you from worrying about emergencies such as earthquakes and fiber cuts. Storage and compute resources can be flexibly configured and elastically scaled as required. This makes resource allocation more accurate and reasonable, greatly improving the resource utilization of big data clusters and reducing the comprehensive analysis cost by 50%.

   In addition, the high performance compute-storage separation architecture breaks the limit of parallel computing of the integrated storage-compute architecture. It maximizes the high bandwidth and high concurrency of OBS, and optimizes the data access efficiency and in-depth parallel computing (such as metadata operation and write algorithm optimization) to improve higher performance.

4. **MRS supports Huawei's in-house Superior Scheduler for better performance.**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS supports Superior Scheduler, a scheduling tool that manages over 10,000 nodes per cluster, enabling larger cluster scales. Superior Scheduler is a scheduling engine designed for the Hadoop YARN distributed resource management system. It is a high-performance and enterprise-level scheduler designed for converged resource pools and multi-tenant service requirements. Superior Scheduler achieves all functions of open-source schedulers, Fair Scheduler, and Capacity Scheduler. Compared with the open-source schedulers, Superior Scheduler is enhanced in the enterprise multi-tenant resource scheduling policy, resource isolation and sharing by multiple users in a tenant, scheduling performance, system resource utilization, and cluster scalability, and is designed to replace open source schedulers.

5. **MRS supports multiple isolation modes and multi-tenant permission management of enterprise-level big data, ensuring higher security.**
   - MRS supports resource deployment and isolation of physical resources in dedicated zones. You can flexibly combine computing and storage resources, such as dedicated computing resources + shared storage resources, shared computing resources + dedicated storage resources, and dedicated computing resources + dedicated storage resources. An MRS cluster supports multiple logical tenants. Permission isolation enables the computing, storage, and table resources of the cluster to be divided based on tenants.
   - With Kerberos authentication, MRS provides role-based access control (RBAC) and sound audit functions.
   - It is proved that with Host Security Service (HSS) interconnected with MRS, service security is enhanced without deteriorating functions and performance.
   - MRS supports unified user login based on web UI. Manager provides user authentication, which grants you permission to access a cluster.
   - MRS supports data storage encryption, encrypted storage of all user accounts and passwords, encrypted transmission of data channels, and bidirectional certificate authentication for cross-trusted-zone data access of service modules.
   - MRS big data clusters provide a complete multi-tenant solution for enterprise-level big data. Multi-tenant refers to a collection of multiple resources (each resource set is a tenant) in an MRS big data cluster. It can allocate and schedule resources, including computing and storage resources. Multi-tenant isolates the resources of a big data cluster into resource sets. Users can lease desired resource sets to run applications and jobs and store data. In a big data cluster, multiple resource sets can be deployed to meet diverse requirements of multiple users.
   - MRS supports enterprise project management. The enterprise project is one way of managing cloud resources. Enterprise Management provides comprehensive management services for enterprise customers, such as cloud resources, personnel, permissions, and financial statuses. Common management consoles are oriented to the control and configuration of individual cloud products. The Enterprise Management console, in contrast, is more focused on resource management. It is designed to help enterprises manage cloud-based resources, personnel, permissions, and finances, in a hierarchical management manner, such as management of companies, departments, and projects. MRS allows users who have enabled Enterprise Project Management Service (EPS) to configure

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

enterprise projects for a cluster during cluster creation and use EPS to manage MRS resources by group. This feature is applicable to scenarios where you need to manage multiple resources by group and perform operations such as permission control and project-based fee query on enterprise projects.

6. **MRS implements HA for all management nodes and supports comprehensive reliability mechanism, making the system more reliable.**

   Based on Apache Hadoop open-source software, MRS optimizes and improves the reliability of main service components.

   - HA for all management nodes

     In the Hadoop open-source version, data and compute nodes are managed in a distributed system, in which a single point of failure (SPOF) does not affect the operation of the entire system. However, a SPOF may occur on management nodes running in centralized mode, which becomes the weakness of the overall system reliability.

     MRS provides similar double-node mechanisms for all management nodes of the service components, such as Manager, Presto, HDFS NameNodes, Hive Servers, HBase HMasters, YARN Resource Managers, Kerberos Servers, and Ldap Servers. All of them are deployed in active/standby mode or configured with load sharing, effectively preventing SPOFs from affecting system reliability.

   - Comprehensive reliability mechanism

     By reliability analysis, the following measures to handle software and hardware exceptions are provided to improve the system reliability:

     - After power supply is restored, services are running properly regardless of a power failure of a single node or the whole cluster, ensuring data reliability in case of unexpected power failures. Key data will not be lost unless the hard disk is damaged.
     - Health status checks and fault handling of the hard disk do not affect services.
     - The file system faults can be automatically handled, and affected services can be automatically restored.
     - The process and node faults can be automatically handled, and affected services can be automatically restored.
     - The network faults can be automatically handled, and affected services can be automatically restored.

7. **MRS provides a visualized big data cluster management interface in a unified manner, making O&M easier.**

   - On the big data cluster management interface, service startup and stopping, configuration modification, and health check are available. MRS also provides visualized and convenient cluster management, monitoring, and alarm functions. Additionally, you can check and audit the system health status in one click, ensuring normal system running and lowering system O&M costs.
   - After Simple Message Notification (SMN) is configured, MRS can send real-time cluster health status information, including cluster changes and

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

component alarms in real time to you through SMS messages or emails, facilitating O&M, real-time monitoring, and real-time alarm sending.

- MRS supports rolling patch upgrade and provides visualized patch release information and one-click patch installation without manual intervention, ensuring long-term stability of user clusters.
- If a problem occurs when you use an MRS cluster, you can initiate O&M authorization on the MRS management console. O&M personnel can help you quickly locate the problem, and you can revoke the authorization at any time. You can also initiate log sharing on the MRS management console to share a specified log scope with O&M personnel, so that O&M personnel can locate faults without accessing the cluster.
- MRS supports to dump logs about cluster creation failures to OBS for O&M personnel to obtain and analyze the logs.

8. **MRS has an open ecosystem and supports seamless interconnection with peripheral services, allowing you to quickly build a unified big data platform.**
   - Based on MRS, a full-stack big data service, enterprises can build a unified big data platform for data ingestion, storage, analysis, and value mining with one click. By interconnecting with DataArts Studio and data visualization services, MRS can help you easily migrate data to the cloud, develop and schedule big data jobs, and display data. This frees you from complex big data platform construction and professional big data calibration and maintenance so you can stay more focused on industry applications and use one copy of data in multiple service scenarios. DataArts Studio is a one-stop data lifecycle development operations platform that provides a broad range of functions, such as data integration, development, governance, service, and visualization. MRS data can be ingested to DataArts Studio for collaborative one-click visualized development by leveraging DataArts Studio's visualized GUI, abundant data development types (script and job), fully-hosted job scheduling and O&M monitoring, and built-in industry data processing pipelines. This makes big data much easier to use, helps you quickly build big data processing centers, and enables fast monetization.
   - MRS is fully compatible with the open source big data ecosystem. With abundant data and application migration tools, MRS helps you quickly migrate data from your own platforms without code modification and service interruption.

1.3 Application Scenarios

Big data is ubiquitous in our lives. MRS is suitable to process big data in the industries such as the Internet of things (IoT), e-commerce, finance, manufacturing, healthcare, energy, and government departments.

Large-scale data analysis

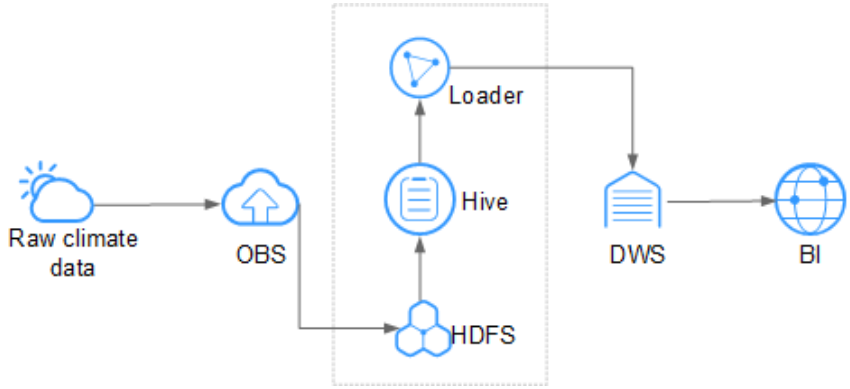
Large-scale data analysis is a major scenario in modern big data systems. Generally, an enterprise has multiple data sources. After data is ingested, extract, transform, and load (ETL) processing is required to generate modelized data for each service module to analyze and sort out. This type of service has the following characteristics:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The requirements for real-time execution are not high, and job execution time ranges from dozens of minutes to hours.
- The data volume is large.
- There are various data sources and diversified formats.
- Data processing usually consists of multiple tasks, and resources need to be planned in detail.

In the environmental protection industry, climate data is stored on OBS and periodically dumped into HDFS for batch analysis. 10 TB of climate data can be analyzed in 1 hour.

Figure 1-2 Large-scale data analysis in the environmental protection industry



MRS has the following advantages in this scenario.

- Low cost: OBS offers cost-effective storage.
- Massive data analysis: TB/PB-level data is analyzed by Hive.
- Visualized data import and export tool: Loader exports data to Data Warehouse Service (DWS) for business intelligence (BI) analysis.

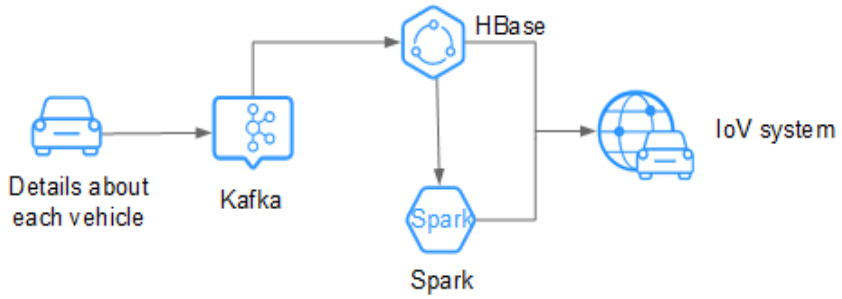
Large-scale data storage

A user who has a large amount of structured data usually requires index-based quasi-real-time query capabilities. For example, in an Internet of Vehicles (IoV) scenario, vehicle maintenance information is queried by vehicle number. Therefore, vehicle information is indexed based on vehicle numbers when it is being stored, to implement second-level response in this scenario. Generally, the data volume is large. The user may store data for one to three years.

For example, in the IoV industry, an automobile company stores data on HBase, which supports PB-level storage and CDR queries in milliseconds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-3 Large-scale data storage in the IoV industry



MRS has the following advantages in this scenario.

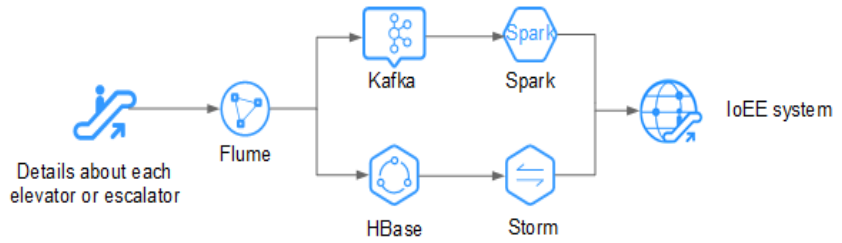
- Real time: Kafka accesses massive amounts of vehicle messages in real time.
- Massive data storage: HBase stores massive volumes of data and supports data queries in milliseconds.
- Distributed data query: Spark analyzes and queries massive volumes of data.

Real-time data processing

Real-time data processing is usually used in scenarios such as anomaly detection, fraud detection, rule-based alarming, and service process monitoring. Data is processed while it is being inputted to the system.

For example, in the Internet of elevators & escalators (IoEE) industry, data of smart elevators and escalators is imported to MRS streaming clusters in real time for real-time alarming.

Figure 1-4 Low-latency streaming processing in the IoEE industry



MRS has the following advantages in this scenario.

- Real-time data ingestion: Flume implements real-time data ingestion and provides various data collection and storage access methods.
- Data source access: Kafka accesses data of tens of thousands of elevators and escalators in real time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.4 MRS Cluster Version Overview

Cluster Version Type

There are two MRS cluster versions: normal and LTS. The components and features of clusters of different versions are slightly different. You can select a version based on service requirements.

- **Normal**
  - **Functions**

    This normal version provides basic cluster operations, such as configuration, management, and O&M.
  - **Components**

    In addition to common components, clusters of the normal version also support components such as Presto, Impala, Kudu, and Sqoop. You can select components based on cluster versions. For details about components of clusters of different versions, see **List of MRS Component Versions**.
- **LTS (long term support)**
  - **Functions**

    In addition to basic cluster operations, clusters of the LTS version support version upgrade.
  - **Components**

    In addition to common components, clusters of the LTS version also support HetuEngine and IoTDB. You can select components based on cluster versions. For details about components of clusters of different versions, see **List of MRS Component Versions**.

Recommendation

- Clusters of the LTS version support version upgrade. To make your clusters upgradable, choose the LTS version.
- Clusters of the LTS version can be deployed in different AZs to implement cross-AZ DR. To make your clusters more secure and have higher DR capabilities, choose the LTS version.
- Clusters of the LTS version support components such as HetuEngine. To use these components, create an LTS version cluster.

  NOTE

  A created LTS cluster cannot be switched to the normal edition.

1.5 List of MRS Component Versions

Components and Versions

Table 1-1 lists the components and their versions required by each MRS cluster version.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- Hadoop components include HDFS, YARN, and MapReduce. DBService, KrbServer, and LdapServer are used in a cluster but not displayed in the component list during cluster creation.
- MRS component versions are consistent with open-source component versions.
- Components in an MRS cluster cannot be upgraded independently. You need to select the cluster version that contains the desired component versions.
- For details about the differences between Long Term Support (LTS) clusters and common clusters, see **MRS Cluster Version Overview**.
- Some cluster versions are restricted. To use them, submit a service ticket.

Table 1-1 MRS component versions

Component	MRS 2.1.0	MRS 3.0.5	MRS 3.1.0	MRS 3.2.0-LTS.1	MRS 3.3.0-LTS	MRS 3.5.0-LTS
Alluxio	-	2.3.0	-	-	-	-
CDL	-	-	-	1.0.0	-	-
ClickHouse	-	20.8.7.15	21.3.4.25	22.3.2.2	23.3.2.37	23.3.2.37
DBService	1.0.0	2.7.0	2.7.0	2.7.0	2.7.0	2.7.0
Doris	-	-	-	-	1.2.3	2.0.13
Flink	1.7.0	1.10.0	1.12.0	1.15.0	1.15.0	1.17.1
Flume	1.6.0	1.9.0	1.9.0	1.9.0	1.11.0	1.11.0
Guardian	-	-	-	-	0.1.0	0.1.0
HBase	2.1.1	2.2.3	2.2.3	2.2.3	2.4.14	2.4.14
HDFS	3.1.1	3.1.1	3.1.1	3.1.1	3.1.1	3.3.1
HetuEngine	-	-	-	1.2.0	2.0.0	2.1.0
Hive	3.1.0	3.1.0	3.1.0	3.1.0	3.1.0	3.1.0
Hue	3.11.0	4.7.0	4.7.0	4.7.0	-	-
Impala	3.2.0	3.4.0	3.4.0	-	-	4.3.0
IoTDB	-	-	-	0.14.0	1.1.0	-
JobGateway	-	-	-	-	1.0.0	1.0.0
Kafka	1.1.0	2.11-2.4.0	2.11-2.4.0	2.11-2.4.0	2.12-2.8.1	2.12-3.6.1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	MRS 2.1.0	MRS 3.0.5	MRS 3.1.0	MRS 3.2.0-LTS.1	MRS 3.3.0-LTS	MRS 3.5.0-LTS
KrbServer	1.15.2	1.17	1.17	1.18	1.20	1.20
Kudu	1.9.0	1.12.1	1.12.1	-	-	1.17.0
LdapServer	1.0.0	2.7.0	2.7.0	2.7.0	2.7.0	2.7.0
Loader	2.0.0	1.99.3	1.99.3	1.99.3	1.99.3	-
MapReduce	3.1.1	3.1.1	3.1.1	3.1.1	3.1.1	3.3.1
MemArtsCC	-	-	-	-	-	3.22
Oozie	-	5.1.0	5.1.0	5.1.0	5.1.0	5.1.0
Presto	308	333	333	-	-	-
Phoenix (integrated in HBase)	-	5.0.0	5.0.0	5.0.0	5.0.0	5.1.2
Ranger	-	2.0.0	2.0.0	2.0.0	2.3.0	2.3.0
Spark	2.3.2	-	-	-	3.3.1	3.3.1
Spark2x	-	2.4.5	2.4.5	3.1.1	-	-
Sqoop	-	-	1.4.7	1.4.7	-	1.4.7
Storm	1.2.1	1.2.1	-	-	-	-
Tez	0.9.1	0.9.2	0.9.2	0.9.2	0.10.2	0.10.2
YARN	3.1.1	3.1.1	3.1.1	3.1.1	3.1.1	3.3.1
ZooKeeper	3.5.1	3.5.6	3.5.6	3.6.3	3.8.1	3.8.1
MRS Manager	2.1.0	-	-	-	-	-
FusionInsight Manager	-	8.0.2.1	8.1.0	8.1.2	8.3.0	8.5.0

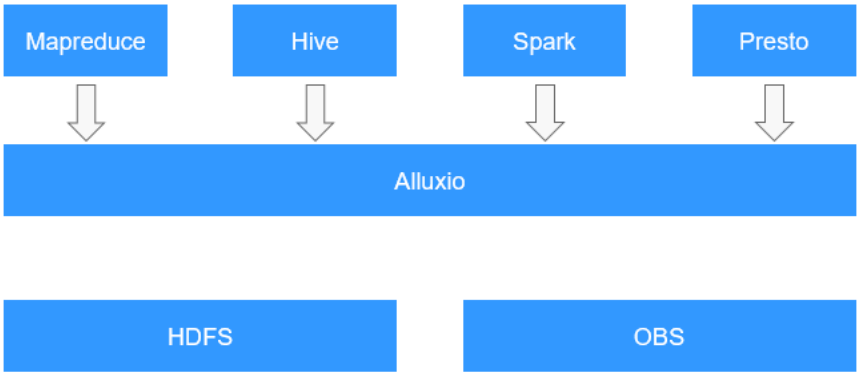
1.6 Components

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.1 Alluxio

Alluxio is data orchestration technology for analytics and AI for the cloud. In the MRS big data ecosystem, Alluxio lies between computing and storage. It provides a data abstraction layer for computing frameworks including Apache Spark, Presto, MapReduce, and Apache Hive, so that upper-layer computing applications can access persistent storage systems including HDFS and OBS through unified client APIs and a global namespace. In this way, computing and storage are separated.

Figure 1-5 Alluxio architecture



Advantages:

- Provides in-memory I/O throughput, and makes elastically scale data-driven applications cost effective.
- Simplified cloud and object storage access
- Simplified data management and a single point of access to multiple data sources
- Easy application deployment

For details about Alluxio, visit **https://docs.alluxio.io/os/user/stable/en/Overview.html**.

1.6.2 ClickHouse

Introduction to ClickHouse

ClickHouse is an open-source columnar database oriented to online analysis and processing. It is independent of the Hadoop big data system and features compression rate and fast query performance. In addition, ClickHouse supports SQL query and provides good query performance, especially the aggregation analysis and query performance based on large and wide tables. The query speed is one order of magnitude faster than that of other analytical databases.

The core functions of ClickHouse are as follows:

Comprehensive DBMS functions

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ClickHouse is a database management system (DBMS) that provides the following basic functions:

- Data Definition Language (DDL): allows databases, tables, and views to be dynamically created, modified, or deleted without restarting services.
- Data Manipulation Language (DML): allows data to be queried, inserted, modified, or deleted dynamically.
- Permission control: supports user-based database or table operation permission settings to ensure data security.
- Data backup and restoration: supports data backup, export, import, and restoration to meet the requirements of the production environment.
- Distributed management: provides the cluster mode to automatically manage multiple database nodes.

Column-based storage and data compression

ClickHouse is a database that uses column-based storage. Data is organized by column. Data in the same column is stored together, and data in different columns is stored in different files.

During data query, columnar storage can reduce the data scanning range and data transmission size, thereby improving data query efficiency.

In a traditional row-based database system, data is stored in the sequence in **Table 1-2**:

Table 1-2 Row-based database

row	ID	Flag	Name	Event	Time
0	12345678901	0	name1	1	2020/1/11 15:19
1	32345678901	1	name2	1	2020/5/12 18:10
2	42345678901	1	name3	1	2020/6/13 17:38
N	...	...	...	...	...

In a row-based database, data in the same row is physically stored together. In a column-based database system, data is stored in the sequence in **Table 1-3**:

Table 1-3 Columnar database

row:	0	1	2	N
ID:	12345678901	32345678901	42345678901	...
Flag:	0	1	1	...
Name:	name1	name2	name3	...

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Event:	1	1	1	...
Time:	2020/1/11 15:19	2020/5/12 18:10	2020/6/13 17:38	...

This example shows only the arrangement of data in a columnar database. Columnar databases store data in the same column together and data in different columns separately. Columnar databases are more suitable for online analytical processing (OLAP) scenarios.

Vectorized executor

ClickHouse uses CPU's Single Instruction Multiple Data (SIMD) to implement vectorized execution. SIMD is an implementation mode that uses a single instruction to operate multiple pieces of data and improves performance with data parallelism (other methods include instruction-level parallelism and thread-level parallelism). The principle of SIMD is to implement parallel data operations at the CPU register level.

Relational model and SQL query

ClickHouse uses SQL as the query language and provides standard SQL query APIs for existing third-party analysis visualization systems to easily integrate with ClickHouse.

In addition, ClickHouse uses a relational model. Therefore, the cost of migrating the system built on a traditional relational database or data warehouse to ClickHouse is lower.

Data sharding and distributed query

The ClickHouse cluster consists of one or more shards, and each shard corresponds to one ClickHouse service node. The maximum number of shards depends on the number of nodes (one shard corresponds to only one service node).

ClickHouse introduces the concepts of local table and distributed table. A local table is equivalent to a data shard. A distributed table itself does not store any data. It is an access proxy of the local table and functions as the sharding middleware. With the help of distributed tables, multiple data shards can be accessed by using the proxy, thereby implementing distributed query.

ClickHouse Applications

ClickHouse is short for Click Stream and Data Warehouse. It is initially applied to a web traffic analysis tool to perform OLAP analysis for data warehouses based on page click event flows. Currently, ClickHouse is widely used in Internet advertising, app and web traffic analysis, telecommunications, finance, and Internet of Things (IoT) fields. It is applicable to business intelligence application scenarios and has a large number of applications and practices worldwide. For details, visit **https://clickhouse.tech/docs/en/introduction/adopters/**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

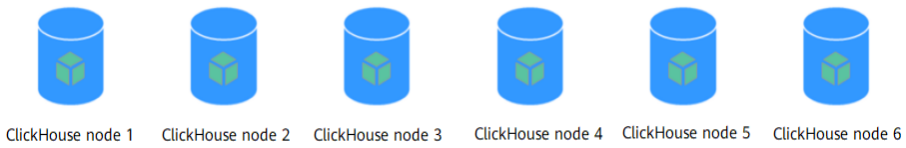
ClickHouse Enhanced Open Source Features

MRS ClickHouse has advantages such as automatic cluster mode, HA deployment, and smooth and elastic scaling.

- Automatic Cluster Mode

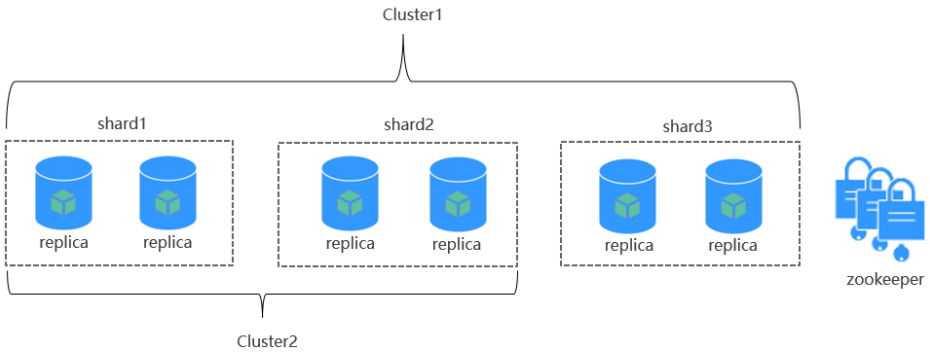
  As shown in **Figure 1-6**, a cluster consists of multiple ClickHouse nodes, which has no central node. It is more of a static resource pool. If the ClickHouse cluster mode is used for services, you need to pre-define the cluster information in the configuration file of each node. Only in this way, services can be correctly accessed.

  Figure 1-6 ClickHouse cluster

  Users are unaware of data partitions and replica storage in common database systems. However, ClickHouse allows you to proactively plan and define detailed configurations such as shards, partitions, and replica locations. The ClickHouse instance of MRS packs the work in a unified manner and adapts it to the automatic mode, implementing unified management, which is flexible and easy to use. A ClickHouse instance consists of three ZooKeeper nodes and multiple ClickHouse nodes. The Dedicated Replica mode is used to ensure high reliability of dual data copies.

  Figure 1-7 ClickHouse cluster structure

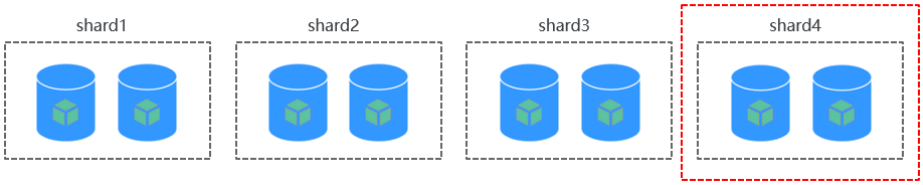
  

- Smooth and Elastic Scaling

  As business grows rapidly, MRS provides ClickHouse, a data migration tool, for scenarios such as the cluster's storage capacity or CPU compute resources approaching the limit. This tool is used to migrate some partitions of one or multiple MergeTree tables on several ClickHouseServer nodes to the same tables on other ClickHouseServer nodes. In this way, service availability is ensured and smooth capacity expansion is implemented.

  When you add ClickHouse nodes to a cluster, use this tool to migrate some data from the existing nodes to the new ones for data balancing after the expansion.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

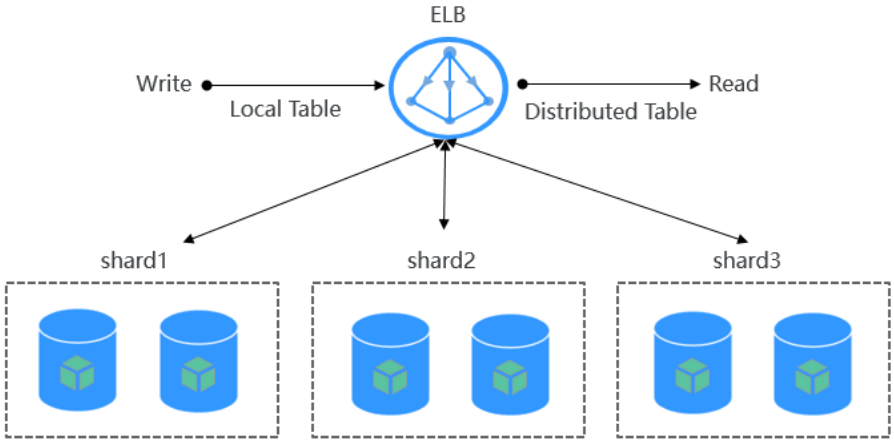


- HA Deployment Architecture

  MRS uses the ELB-based high availability (HA) deployment architecture to automatically distribute user access traffic to multiple backend nodes, expanding service capabilities to external systems and improving fault tolerance. As shown in **Figure 1-8**, when a client application requests a cluster, Elastic Load Balance (ELB) is used to distribute traffic. With the ELB polling mechanism, data is written to local tables and read from distributed tables on different nodes. In this way, data read/write load and high availability of application access are guaranteed.

  After the ClickHouse cluster is provisioned, each ClickHouse instance node in the cluster corresponds to a replica, and two replicas form a logical shard. For example, when creating a ReplicatedMergeTree table, you can specify shards so that data can be automatically synchronized between two replicas in the same shard.

  Figure 1-8 HA deployment architecture

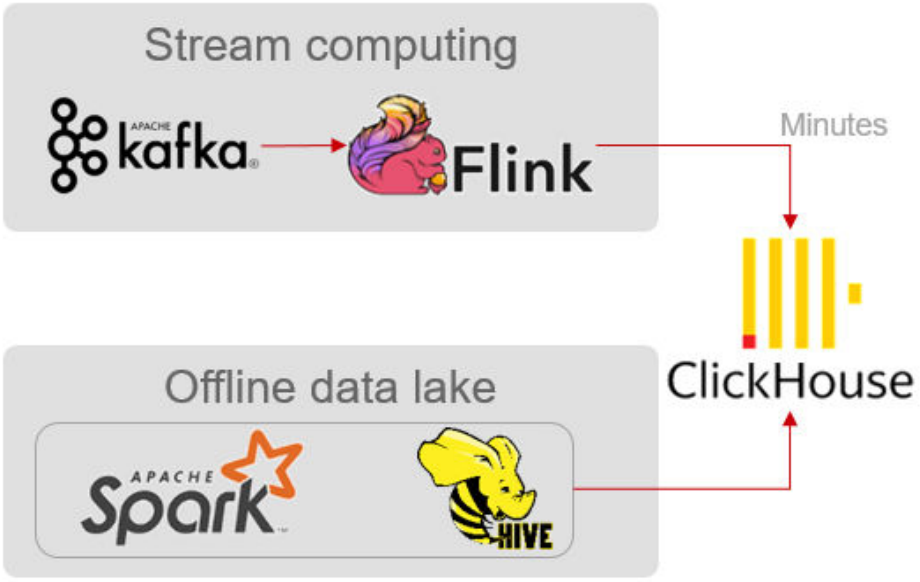


Relationships Between ClickHouse and Other Components

ClickHouse depends on ZooKeeper for installation and deployment.

Flink stream computing applications are used to generate common report data (detail flat-wide tables) and write the report data to ClickHouse in quasi-real time. Hive/Spark jobs are used to generate common report data (detail flat-wide tables) and batch import the data to ClickHouse.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



NOTE

Currently, ClickHouse does not support interconnection with Kafka in normal mode or HDFS in security mode.

1.6.3 CDL

Overview

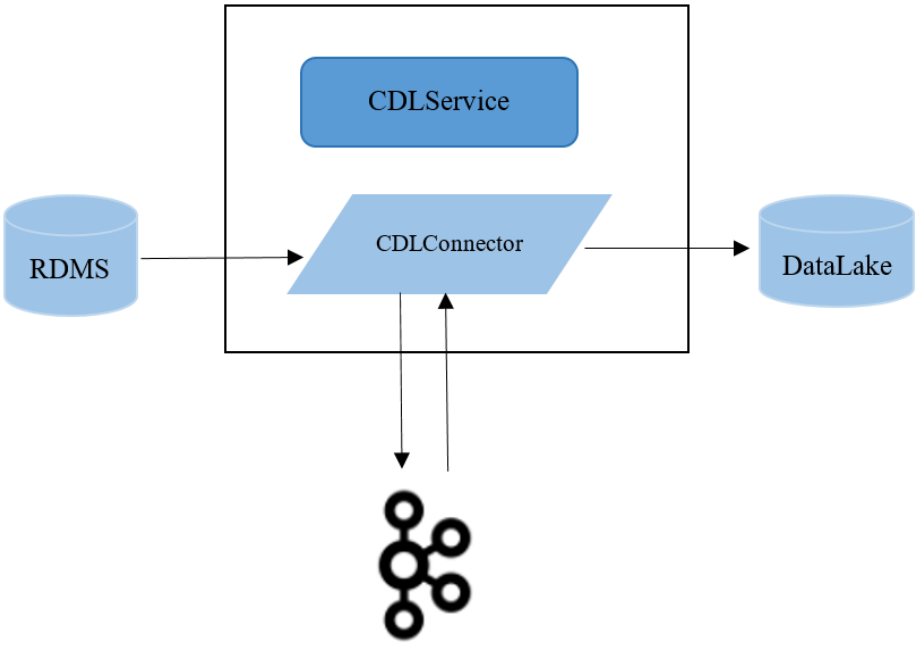
Change Data Loader (CDL) is a real-time data integration service based on Kafka Connect. The CDL service captures data change events from various OLTP databases and pushes them to Kafka. Then, Sink Connector pushes the events to the big data ecosystem.

Currently, CDL supports MySQL, PostgreSQL, Hudi, Kafka, and ThirdParty-Kafka data sources. Data can be written to Kafka, Hudi, GaussDB(DWS), and ClickHouse.

CDL structure

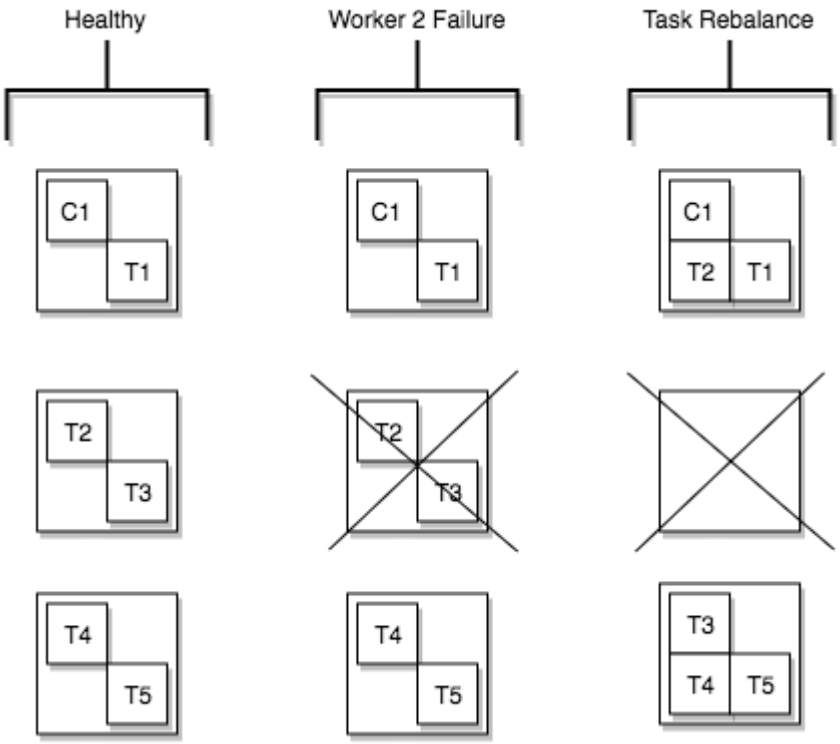
The CDL service has two important roles: CDLConnector and CDLService. CDLConnector, including Source Connector and Sink Connector, executes data capture jobs. CDLService manages and creates jobs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



The CDLService instances of the CDL service work in multi-active mode. Any CDLService instance can perform service operations. The CDLConnector instances work in distributed mode and provide HA and rebalance capabilities. When tasks are created, the number of tasks specified is balanced among CDLConnector instances in a cluster to ensure that the number of tasks running on each instance is similar. If a CDLConnector instance is abnormal or a node breaks down, the number of tasks are rebalanced on other nodes.

Figure 1-9 Rebalance of a task



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Relationship Between CDL and Other Components

The CDL component is based on the Kafka Connect framework. Captured data is forwarded using Kafka topics. Therefore, the CDL component depends on the Kafka component. In addition, the CDL component stores task metadata and monitoring information that are also stored in a database. Therefore, the CDL component also depends on the DBService component.

1.6.4 DBService

Overview

DBService is a HA storage system for relational databases, which is applicable to the scenario where a small amount of data (about 10 GB) needs to be stored, for example, component metadata. DBService can only be used by internal components of a cluster and provides data storage, query, and deletion functions.

DBService is a basic component of a cluster. Components such as Hive, Hue, Oozie, Loader, CDL, Flink, HetuEngine, Kafka, Metadata, and Ranger store their metadata in DBService, and provide the metadata backup and restoration functions using DBService.

DBService Architecture

DBService in the cluster works in active/standby mode. Two DBServer instances are deployed and each instance contains three modules: HA, Database, and FloatIP.

Figure 1-10 shows the DBService logical architecture.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-10 DBService architecture

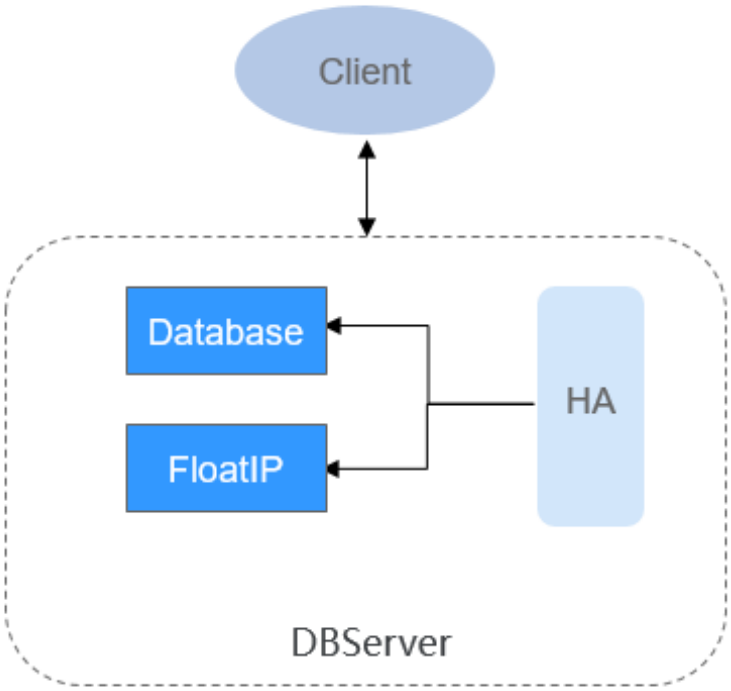


Table 1-4 describes the modules shown in **Figure 1-10**

Table 1-4 Module description

Name	Description
HA	HA management module. The active/standby DBServer uses the HA module for management.
Database	Database module. This module stores the metadata of the Client module.
FloatIP	Floating IP address that provides the access function externally. It is enabled only on the active DBServer instance and is used by the Client module to access Database.
Client	Client using the DBService component, which is deployed on the component instance node. The client connects to the database by using FloatIP and then performs metadata adding, deleting, and modifying operations.

Relationship Between DBService and Other Components

DBService is a basic component of a cluster. Components such as Hive, Hue, Oozie, and Loader store their metadata in DBService, and provide the metadata backup and restoration functions by using DBService.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.5 Doris

1.6.5.1 Basic Principles

Introduction to Doris

Doris is a high-performance, real-time analytical database based on the MPP architecture, known for its extreme speed and ease of use. It returns query results of large-scale data in sub-seconds and supports high-concurrency point queries and high-throughput complex analysis. All this makes Doris an ideal tool for report analysis, ad-hoc query, unified data warehouse, and data lake query acceleration. On Doris, users can build various applications, such as user behavior analysis, AB test platform, log retrieval analysis, user portrait analysis, and order analysis. For more information, see **Apache Doris**.

Doris Architecture

The following figure shows the overall architecture of Doris. Both frontend (FE) and backend (BE) nodes are horizontally scalable.

Figure 1-11 Doris architecture

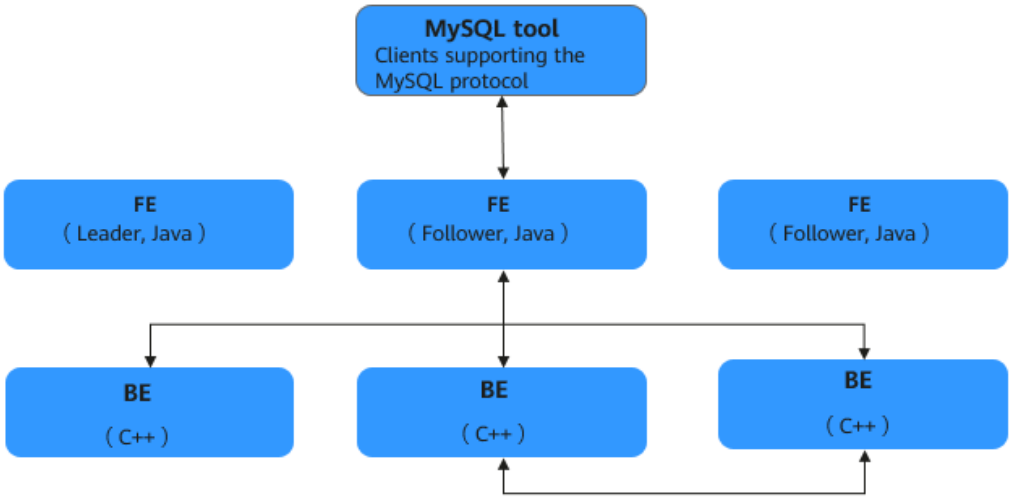


Table 1-5 Description

Parameter	Description
MySQL Tools	Doris is fully compatible with MySQL dialect and can be accessed by various client tools. It also supports standard SQL and can seamlessly connect to BI tools.
FE	Frontend nodes process user access requests, plan query parsing, and manage metadata and nodes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
BE	Backend nodes store data, execute query plans, and balance load among copies.
Leader	Leader is a role elected from Follower nodes.
Follower	Follower nodes receive metadata logs, which must be written successfully in most nodes.

Doris uses the MPP model for inter-node and intra-node parallel execution, making it suitable for distributed joins of large tables.

It also supports vectorized query execution engines, adaptive query execution (AQE) technology, optimization strategies that combine CBO and RBO, and hot data cache queries.

Basic Concepts

In Doris, data is logically described in the form of tables.

- **Rows and Columns**

  A table consists of rows and columns.

  - Row: a row of user data.
  - Column: different fields in a row of data.

  Columns can be classified into two types: keys and values. From the service perspective, Key and Value correspond to dimension columns and metric columns, respectively. In the aggregation model, rows with the same Key column are aggregated into one row. How Value columns are aggregated as specified by a user when the table is created.

- **Tablets and Partitions**

  In the Doris storage engine, user data is horizontally divided into several tablets (also called data buckets). Each tablet contains several rows of data. The data between the individual tablets does not intersect and is physically stored independently.

  Multiple tablets logically belong to different partitions. A tablet belongs to only one partition, but a partition can contain multiple tablets. Since the tablets are physically stored independently, the partitions can be seen as physically independent, too. Tablet is the smallest physical storage unit for data operations such as movement and replication.

  Multiple partitions form a table. A partition can be regarded as the smallest logical unit for management. Data can be imported or deleted only for one partition.

- **Data Models**

  Doris data models are classified into three types: Aggregate, Unique, and Duplicate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Aggregate Model**

  When data is imported, rows with the same Key column are aggregated, and the Value columns are aggregated based on the AggregationType configured by users. AggregationType has the following modes:

  - SUM: Sum up the values in multiple rows.
  - REPLACE: Replace the previous value with the newly imported value.
  - MAX: Keep the maximum value.
  - MIN: Keep the minimum value.

- **Unique Model**

  In some multi-dimensional analysis scenarios, users are highly concerned about how to create uniqueness constraints for the Primary Key. The Unique model is introduced to solve this problem.

  - **Merge on Read**

    The merge on read implementation in the Unique model is equivalent to Replace implementation in the Aggregate model. The internal implementation and data storage method are the same.

  - **Merge on Write**

    The Merge on Write implementation of the Unique model is completely different from that of the Aggregate model. It can deliver better performance (almost like that of the Duplicate model) in aggregation queries with primary key limitations. This implementation is particularly suitable for aggregation queries and those using indexes to filter out large scale data.

    In a Unique table where Merge on Write is enabled, overwritten and updated data is marked and deleted during data import, and new data is written to a new file. During a query, all data marked for deletion is filtered out at the file level, and the read data is the latest data. This eliminates the data aggregation process in Merge on Read and supports pushdown of multiple predicates in many cases. Performance can be greatly improved in many scenarios, especially in the case of aggregation queries.

- **Duplicate Model**

  In some multi-dimensional analysis scenarios, primary keys and data aggregation are not required. Duplicate models can be introduced to meet such requirements.

  Different from the Aggregate and Unique models, the Duplicate model stores the data as it is and executes no aggregation. Even if there are two identical rows of data, they will both be retained. The **DUPLICATE KEY** in the **CREATE TABLE** statement is only used to specify based on which columns the data are sorted.

- **Data Model Selection**

  The data model is established when the table is created and cannot be modified. Therefore, it is important to select a proper data model.

  - The Aggregate model aggregates data in advance, greatly reducing data scanning and calculation workload. Therefore, it is suitable for

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

reporting query business, which has fixed schema. However, this model is not user-friendly for **count(*)** queries. Since the aggregation method on the Value column is fixed, semantic correctness should be considered in other types of aggregation queries.

- The Unique model ensures that the primary key is unique when it is required. However, pre-aggregation such as Rollup cannot be used in this case.

  The Unique model supports only the update of an entire row. If you need to update both the unique primary key constraint and some columns (for example, importing multiple source tables to one Doris table), you can use the Aggregate model and set the aggregation type of non-primary key columns to **REPLACE_IF_NOT_NULL**.

- Duplicate is suitable for ad-hoc queries in any dimension. Although pre-aggregation cannot be used, Duplicate is not restricted by the aggregation model and can make full use of the advantages of the column-store model, that is, only related columns are read, and not all key columns need to be read.

1.6.5.2 Relationship with Other Components

Relationship Between Doris and HDFS

Doris can import and export HDFS data and directly query HDFS data sources.

Relationship Between Doris and Hudi

Doris can directly query Hudi data sources.

Relationship Between Doris and Spark

Spark Doris Connector allows Spark to read data stored in Doris and write data to Doris.

Relationship Between Doris and Flink

Flink Doris Connector allows you to perform operations (read, insert, modify, and delete) on data stored in Doris through Flink.

Relationship Between Doris and Hive

Doris can directly query Hive data sources.

Relationship Between Doris and Kafka

Doris can import Kafka data.

1.6.6 Flink

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.6.1 Flink Basic Principles

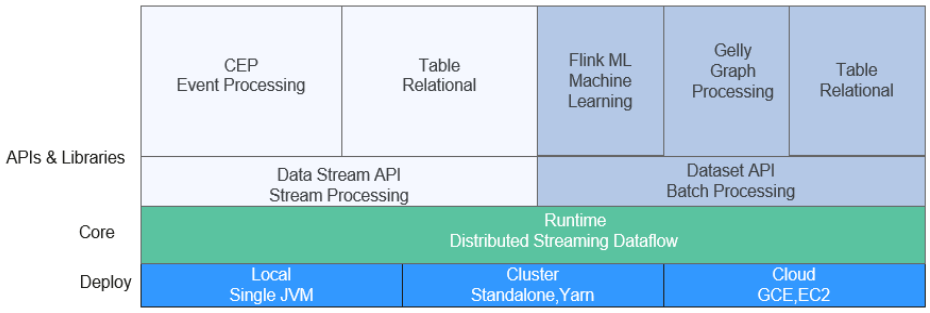
Overview

Flink is a unified computing framework that supports both batch processing and stream processing. It provides a stream data processing engine that supports data distribution and parallel computing. Flink features stream processing and is a top open source stream processing engine in the industry.

Flink provides high-concurrency pipeline data processing, millisecond-level latency, and high reliability, making it extremely suitable for low-latency data processing.

Figure 1-12 shows the technology stack of Flink.

Figure 1-12 Technology stack of Flink



Flink provides the following features in the current version:

- DataStream
- Checkpoint
- Window
- Job Pipeline
- Configuration Table

Flink Architecture

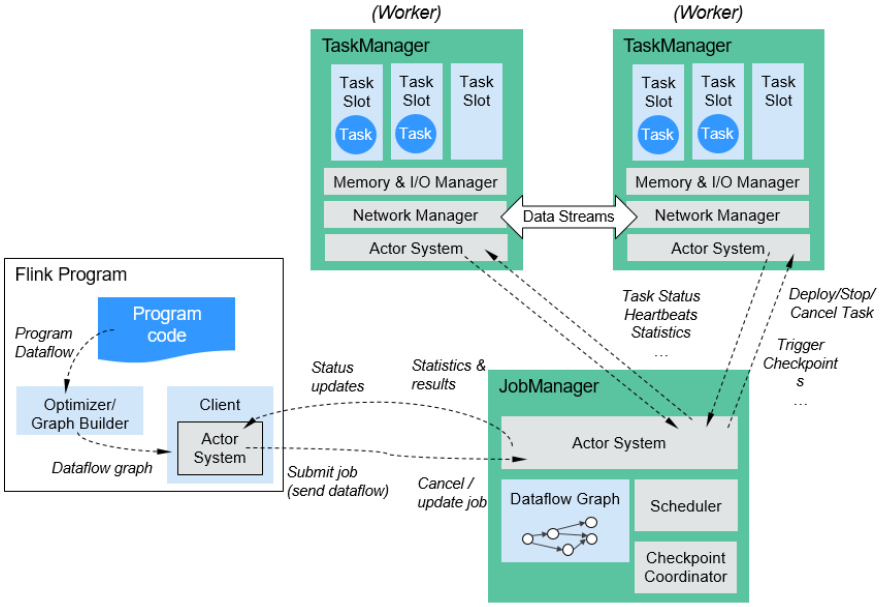
The Flink service has two important roles: FlinkResource and FlinkServer.

- FlinkResource is a mandatory tool that must be installed. It manages client configurations. The role includes the original **lib** package and configuration file used for downloading the client, and the original **lib** package on which FlinkServer depends for job submission. There is no entity process. The job running does not depend on FlinkResource.
- FlinkServer is a web-based secondary development platform for job management. You can develop and manage FlinkSQL jobs on the UI. It features UI-based O&M management and standardized SQL statements for job development.

Figure 1-13 shows the Flink architecture.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-13 Flink architecture



As shown in the above figure, the entire Flink system consists of three parts:

- Client

  Flink client is used to submit jobs (streaming jobs) to Flink.

- TaskManager

  TaskManager is a service execution node of Flink. It executes specific tasks. A Flink system can have multiple TaskManagers. These TaskManagers are equivalent to each other.

- JobManager

  JobManager is a management node of Flink. It manages all TaskManagers and schedules tasks submitted by users to specific TaskManagers. In high-availability (HA) mode, multiple JobManagers are deployed. Among these JobManagers, one is selected as the active JobManager, and the others are standby.

Flink Principles

- **Stream & Transformation & Operator**

  A Flink program consists of two building blocks: stream and transformation.

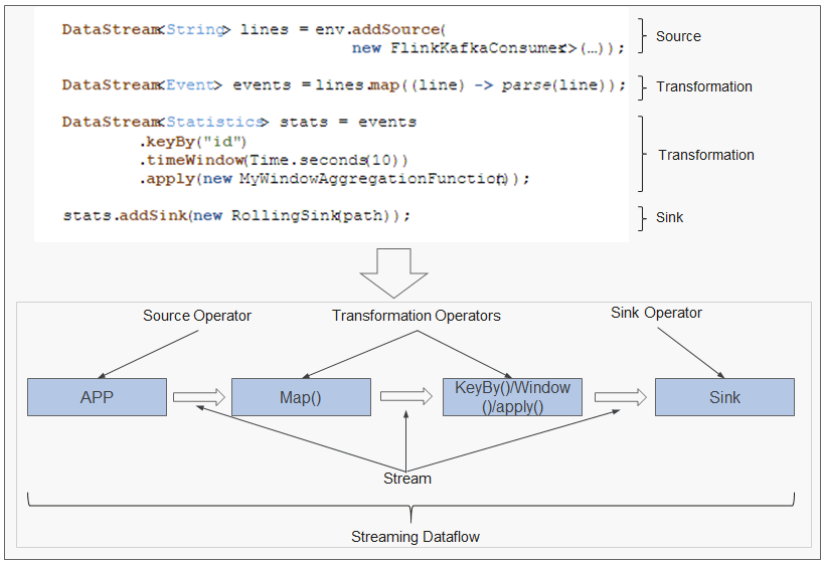
  a. Conceptually, a stream is a (potentially never-ending) flow of data records, and a transformation is an operation that takes one or more streams as input, and produces one or more output streams as a result.

  b. When a Flink program is executed, it is mapped to a streaming dataflow. A streaming dataflow consists of a group of streams and transformation operators. Each dataflow starts with one or more source operators and ends in one or more sink operators. A dataflow resembles a directed acyclic graph (DAG).

  Figure 1-14 shows the streaming dataflow to which a Flink program is mapped.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-14 Example of Flink DataStream



As shown in **Figure 1-14**, **FlinkKafkaConsumer** is a source operator; Map, KeyBy, TimeWindow, and Apply are transformation operators; RollingSink is a sink operator.

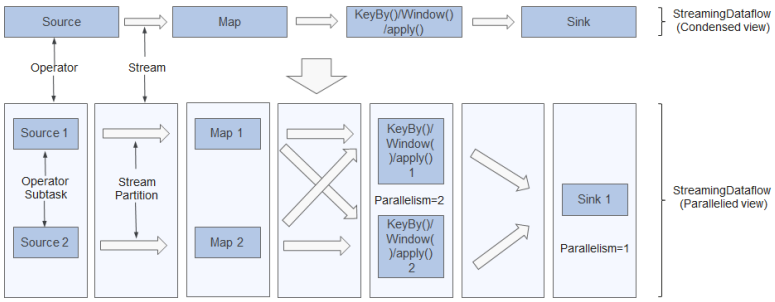
- **Pipeline Dataflow**

  Applications in Flink can be executed in parallel or distributed modes. A stream can be divided into one or more stream partitions, and an operator can be divided into multiple operator subtasks.

  The executor of streams and operators are automatically optimized based on the density of upstream and downstream operators.

  – Operators with low density cannot be optimized. Each operator subtask is separately executed in different threads. The number of operator subtasks is the parallelism of that particular operator. The parallelism (the total number of partitions) of a stream is that of its producing operator. Different operators of the same program may have different levels of parallelism, as shown in **Figure 1-15**.

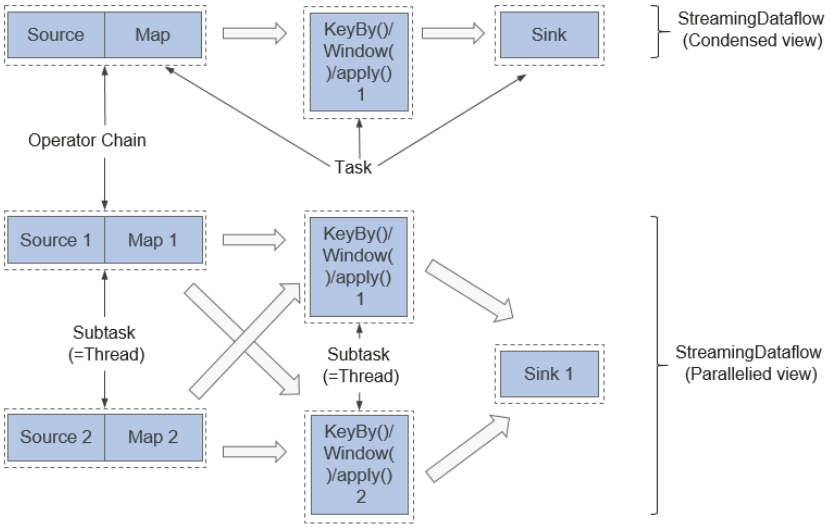
  Figure 1-15 Operator

  – Operators with high density can be optimized. Flink chains operator subtasks together into a task, that is, an operator chain. Each operator chain is executed by one thread on TaskManager, as shown in **Figure 1-16**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-16 Operator chain



- In the upper part of **Figure 1-16**, the condensed Source and Map operators are chained into an Operator Chain, that is, a larger operator. The Operator Chain, KeyBy, and Sink all represent an operator respectively and are connected with each other through streams. Each operator corresponds to one task during the running. Namely, there are three tasks in the upper part.
- In the lower part of **Figure 1-16**, each task, except Sink, is paralleled into two subtasks. The parallelism of the Sink operator is one.

Key Features

- Stream processing

  The real-time stream processing engine features high throughput, high performance, and low latency, which can provide processing capability within milliseconds.

- Various status management

  The stream processing application needs to store the received events or intermediate result in a certain period of time for subsequent access and processing at a certain time point. Flink provides diverse features for status management, including:

  – Multiple basic status types: Flink provides various states for data structures, such as ValueState, ListState, and MapState. Users can select the most efficient and suitable status type based on the service model.

  – Rich State Backend: State Backend manages the status of applications and performs Checkpoint operations as required. Flink provides different State Backends. State can be stored in the memory or RocksDB, and supports the asynchronous and incremental Checkpoint mechanism.

  – Exactly-once state consistency: The Checkpoint and fault recovery capabilities of Flink ensure that the application status of tasks is consistent before and after a fault occurs. Flink supports transactional

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

output for some specific storage devices. In this way, exactly-once output can be ensured even when a fault occurs.

- Various time semantics

  Time is an important part of stream processing applications. For real-time stream processing applications, operations such as window aggregation, detection, and matching based on time semantics are quite common. Flink provides various time semantics.

  - Event-time: The timestamp provided by the event is used for calculation, making it easier to process the events that arrive at a random sequence or arrive late.
  - Watermark: Flink introduces the concept of Watermark to measure the development of event time. Watermark also provides flexible assurance for balancing processing latency and data integrity. When processing event streams with Watermark, Flink provides multiple processing options if data arrives after the calculation, for example, redirecting data (side output) or updating the calculation result.
  - Processing-time and Ingestion-time are supported.
  - Highly flexible streaming window: Flink supports the time window, count window, session window, and data-driven customized window. You can customize the triggering conditions to implement the complex streaming calculation mode.

- Fault tolerance mechanism

  In a distributed system, if a single task or node breaks down or is faulty, the entire task may fail. Flink provides a task-level fault tolerance mechanism, which ensures that user data is not lost when an exception occurs in a task and can be automatically restored.

  - Checkpoint: Flink implements fault tolerance based on checkpoint. Users can customize the checkpoint policy for the entire task. When a task fails, the task can be restored to the status of the latest checkpoint and data after the snapshot is resent from the data source.
  - Savepoint: A savepoint is a consistent snapshot of application status. The savepoint mechanism is similar to that of checkpoint. However, the savepoint mechanism needs to be manually triggered. The savepoint mechanism ensures that the status information of the current stream application is not lost during task upgrade or migration, facilitating task suspension and recovery at any time point.

- Flink SQL

  Table APIs and SQL use Apache Calcite to parse, verify, and optimize queries. Table APIs and SQL can be seamlessly integrated with DataStream and DataSet APIs, and support user-defined scalar functions, aggregation functions, and table value functions. The definition of applications such as data analysis and ETL is simplified. The following code example shows how to use Flink SQL statements to define a counting application that records session times.

  ```
  SELECT userId, COUNT(*)
  FROM clicks
  GROUP BY SESSION(clicktime, INTERVAL '30' MINUTE), userId
  ```

- CEP in SQL

  Flink allows users to represent complex event processing (CEP) query results in SQL for pattern matching and evaluate event streams on Flink.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

CEP SQL is implemented through the **MATCH_RECOGNIZE** SQL syntax. The **MATCH_RECOGNIZE** clause is supported by Oracle SQL since Oracle Database 12c and is used to indicate event pattern matching in SQL. The following is an example of CEP SQL:

```
SELECT T.aid, T.bid, T.cid
FROM MyTable
   MATCH_RECOGNIZE (
     PARTITION BY userid
     ORDER BY proctime
     MEASURES
       A.id AS aid,
       B.id AS bid,
       C.id AS cid
     PATTERN (A B C)
     DEFINE
       A AS name = 'a',
       B AS name = 'b',
       C AS name = 'c'
   ) AS T
```

1.6.6.2 Flink HA Solution

Flink HA Solution

A Flink cluster has only one JobManager. This has the risks of single point of failures (SPOFs). There are three modes of Flink: Flink On Yarn, Flink Standalone, and Flink Local. Flink On Yarn and Flink Standalone modes are based on clusters and Flink Local mode is based on a single node. Flink On Yarn and Flink Standalone provide an HA mechanism. With such a mechanism, you can recover the JobManager from failures and thereby eliminate SPOF risks. This section describes the HA mechanism of the Flink On Yarn.

Flink supports the HA mode and job exception recovery that highly depend on ZooKeeper. If you want to enable the two functions, configure ZooKeeper in the **flink-conf.yaml** file in advance as follows:

```
high-availability: zookeeper
high-availability.zookeeper.quorum:  ZooKeeper IP address:2181
high-availability.storageDir: hdfs:///flink/recovery
```

Yarn

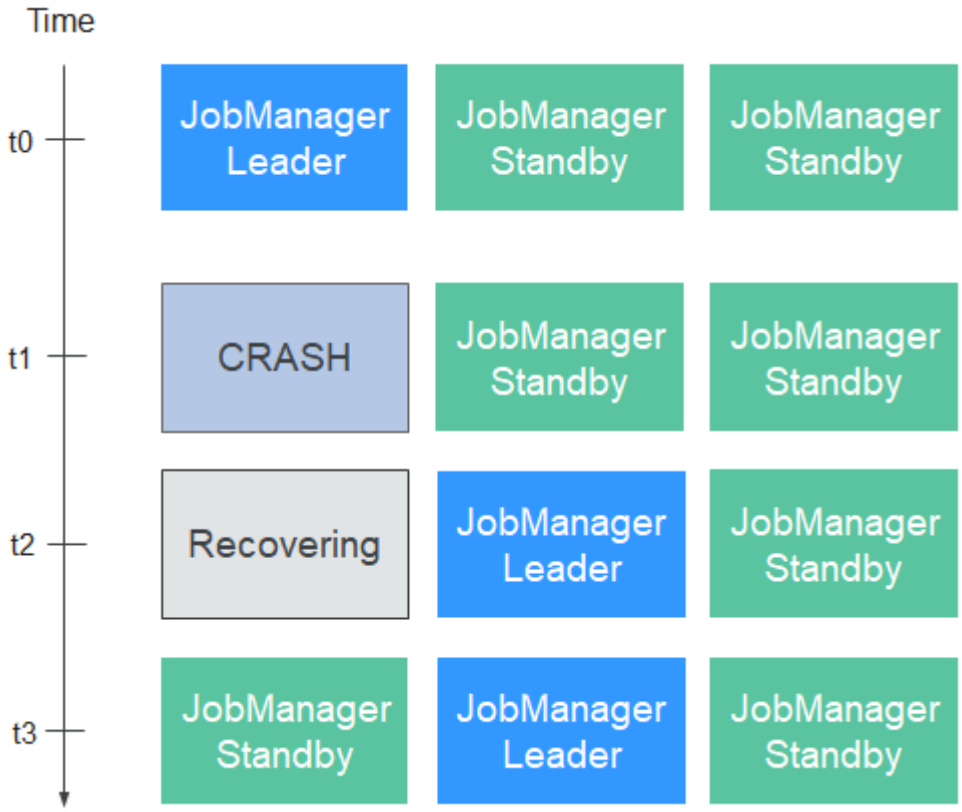
Flink JobManager and Yarn ApplicationMaster are in the same process. Yarn ResourceManager monitors ApplicationMaster. If ApplicationMaster is abnormal, Yarn restarts it and restores all JobManager metadata from HDFS. During the recovery, existing tasks cannot run and new tasks cannot be submitted. ZooKeeper stores JobManager metadata, such as information about jobs, to be used by the new JobManager. A TaskManager failure is listened and processed by the DeathWatch mechanism of Akka on JobManager. When a TaskManager fails, a container is requested again from Yarn and a TaskManager is created.

Standalone

In the standalone mode, multiple JobManagers can be started and ZooKeeper elects one as the Leader JobManager. In this mode, there is a leader JobManager and multiple standby JobManagers. If the leader JobManager fails, a standby JobManager takes over the leadership. **Figure 1-17** shows the process of a leader/standby JobManager switchover.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-17 Switchover process



Restoring TaskManager

A TaskManager failure is listened and processed by the DeathWatch mechanism of Akka on JobManager. If the TaskManager fails, the JobManager creates a TaskManager and migrates services to the created TaskManager.

Restoring JobManager

Flink JobManager and Yarn ApplicationMaster are in the same process. Yarn ResourceManager monitors ApplicationMaster. If ApplicationMaster is abnormal, Yarn restarts it and restores all JobManager metadata from HDFS. During the recovery, existing tasks cannot run and new tasks cannot be submitted.

Restoring Jobs

If you want to restore jobs, ensure that the startup policy is configured in Flink configuration files. Supported restart policies are **fixed-delay**, **failure-rate**, and **none**. Jobs can be restored only when the policy is configured to **fixed-delay** or **failure-rate**. If the restart policy is configured to **none** and **Checkpoint** is configured for **Job**, the restart policy is automatically configured to **fixed-delay** and the value of **restart-strategy.fixed-delay.attempts** specifies the number of retry times.

The configuration strategies are as follows:

```
restart-strategy: fixed-delay
restart-strategy.fixed-delay.attempts: 3
restart-strategy.fixed-delay.delay: 10 s
```

Jobs will be restored in the following scenarios:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If a JobManager fails, all its jobs are stopped, and will be recovered after another JobManager is created and running.
- If a TaskManager fails, all tasks on the TaskManager are stopped, and will be started until there are available resources.
- When a task of a job fails, the job is restarted.

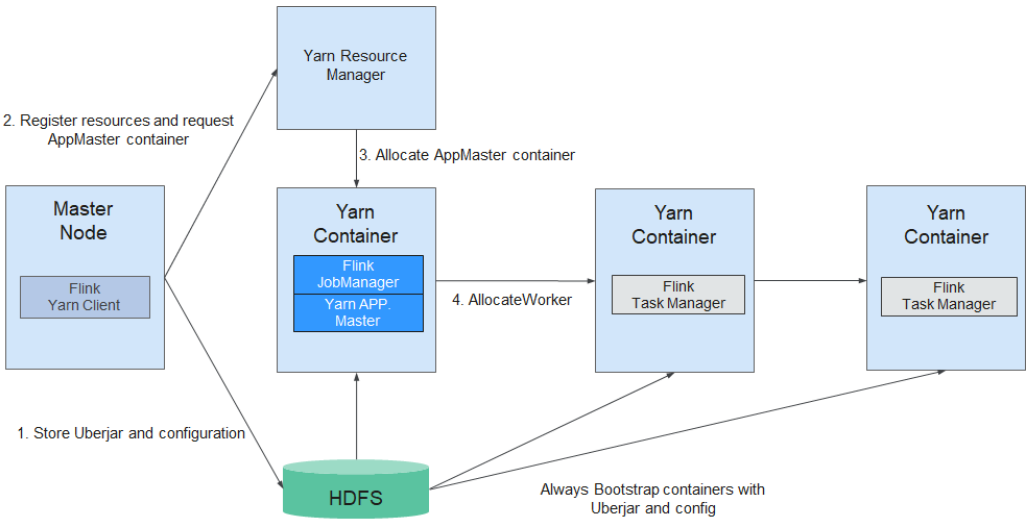
1.6.6.3 Relationships Between Flink and Other Components

Relationship Between Flink and Yarn

Flink supports Yarn-based cluster management mode. In this mode, Flink serves as an application of Yarn and runs on Yarn.

Figure 1-18 shows the Yarn-based Flink cluster deployment.

Figure 1-18 Yarn-based Flink cluster deployment



1. The Flink Yarn Client first checks whether there are sufficient resources for starting the Yarn cluster. If yes, the Flink Yarn client uploads JAR files and configuration files to HDFS.
2. Flink Yarn client communicates with Yarn ResourceManager to request a container for starting ApplicationMaster. After all Yarn NodeManagers finish downloading the JAR file and configuration files, the ApplicationMaster is started.
3. During the startup, the ApplicationMaster interacts with the Yarn ResourceManager to request the container for starting a TaskManager. After the container is ready, the TaskManager process is started.
4. In the Flink Yarn cluster, the ApplicationMaster and Flink JobManager are running in the same container. The ApplicationMaster informs each TaskManager of the RPC address of the JobManager. After TaskManagers are started, they register with the JobManager.
5. After all TaskManagers has registered with the JobManager, Flink starts up in the Yarn cluster. Then, the Flink Yarn client can submit Flink jobs to the JobManager, and Flink can perform mapping, scheduling, and computing for the jobs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.6.4 Flink Sliding Window Enhancement

This section describes the sliding windows of Flink and provides the sliding window optimization method.

Introduction to Window

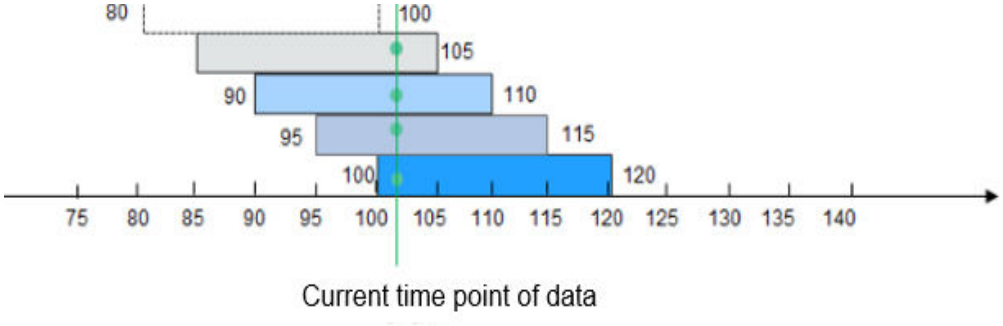
Data in a window is saved as intermediate results or original data. If you perform a sum operation (**window(SlidingEventTimeWindows.of(Time.seconds(20), Time.seconds(5))).sum**) on data in the window, only the intermediate result will be retained. If a custom window (**window(SlidingEventTimeWindows.of(Time.seconds(20), Time.seconds(5))).apply(new UDF)**) is used, all original data in the window will be saved.

If custom windows **SlidingEventTimeWindow** and **SlidingProcessingTimeWindow** are used, data is saved as multiple backups. Assume that the window is defined as follows:

```
window(SlidingEventTimeWindows.of(Time.seconds(20), Time.seconds(5))).apply(new UDFWindowFunction)
```

If a block of data arrives, it is assigned to four different windows (20/5 = 4). That is, the data is saved as four copies in the memory. When the window size or sliding period is set to a large value, data will be saved as excessive copies, causing redundancy.

Figure 1-19 Original structure of a window



If a data block arrives at the 102nd second, it is assigned to windows [85, 105), [90, 110), [95, 115), and [100, 120).

Window Optimization

As mentioned in the preceding, there are excessive data copies when original data is saved in SlidingEventTimeWindow and SlidingProcessingTimeWindow. To resolve this problem, the window that stores the original data is restructured, which optimizes the storage and greatly lowers the storage space. The window optimization scheme is as follows:

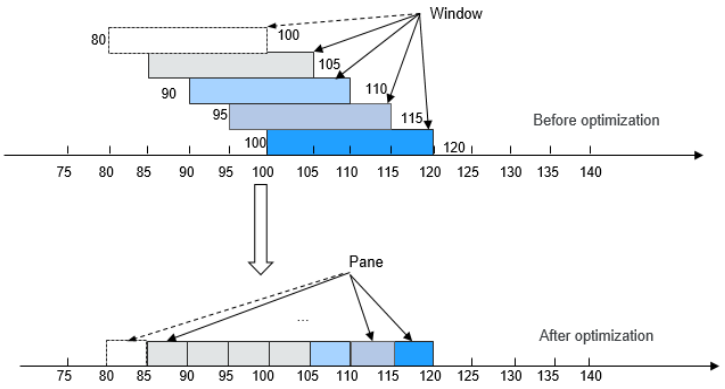
1. Use the sliding period as a unit to divide a window into different panes.

   A window consists of one or multiple panes. A pane is essentially a sliding period. For example, the sliding period (namely, the pane) of **window(SlidingEventTimeWindows.of(Time.seconds(20), Time.seconds.of(5)))** lasts for 5 seconds. If this window ranges from [100,

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

120), this window can be divided into panes [100, 105), [105, 110), [110, 115), and [115, 120).

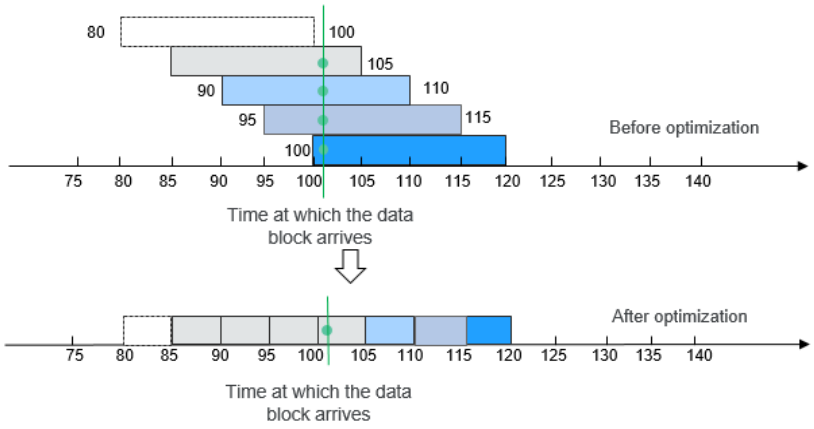
Figure 1-20 Window optimization



2. When a data block arrives, it is not assigned to a specific window. Instead, Flink determines the pane to which the data block belongs based on the timestamp of the data block, and saves the data block into the pane.

   A data block is saved only in one pane. In this case, only a data copy exists in the memory.

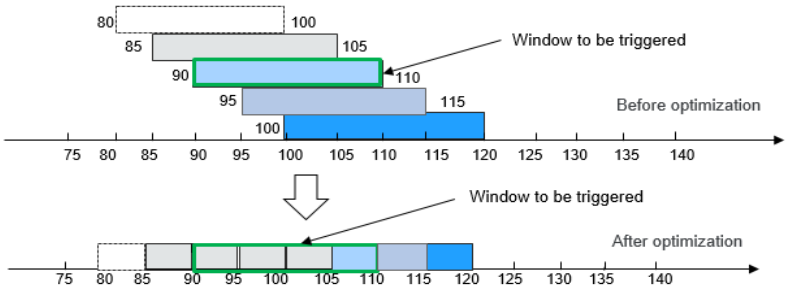
Figure 1-21 Saving data in a window



3. To trigger a window, compute all panes contained in the window, and combine all these panes into a complete window.

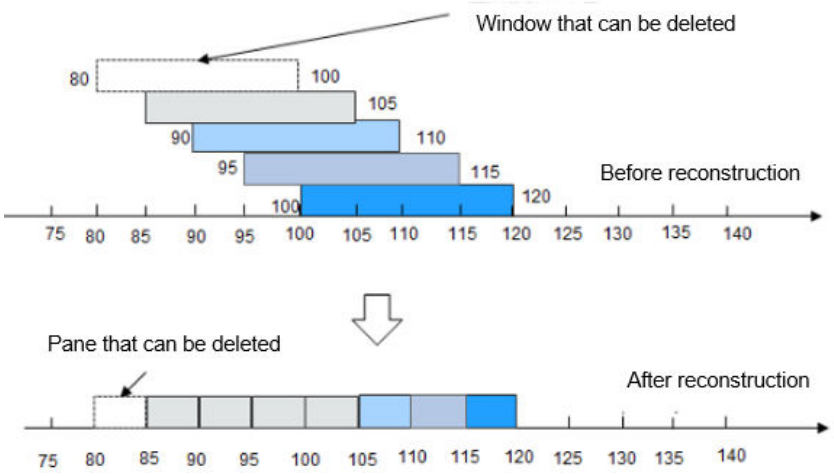
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-22 Triggering a window



4. If a pane is not required, you can delete it from the memory.

Figure 1-23 Deleting a window



After optimization, the quantity of data copies in the memory and snapshot is greatly reduced.

1.6.6.5 Flink Job Pipeline Enhancement

Generally, logic code related to a service is stored in a large JAR package, which is called Fat JAR.

Disadvantages of Fat JAR are as follows:

- When service logic becomes more and more complex, the size of the Fat JAR increases.
- Fat Jar makes coordination complex. Developers of all services are working with the same service logic. Even though the service logic can be divided into several modules, all modules are tightly coupled with each other. If the requirement needs to be changed, the entire flow diagram needs to be replanned.

Splitting of jobs is facing the following problems:

- Data transmission between jobs can be achieved using Kafka. For example, job A transmits data to the topic A in Kafka, and then job B and job C read data from the topic A in Kafka. This solution is simple and easy to implement, but the latency is always longer than 100 ms.

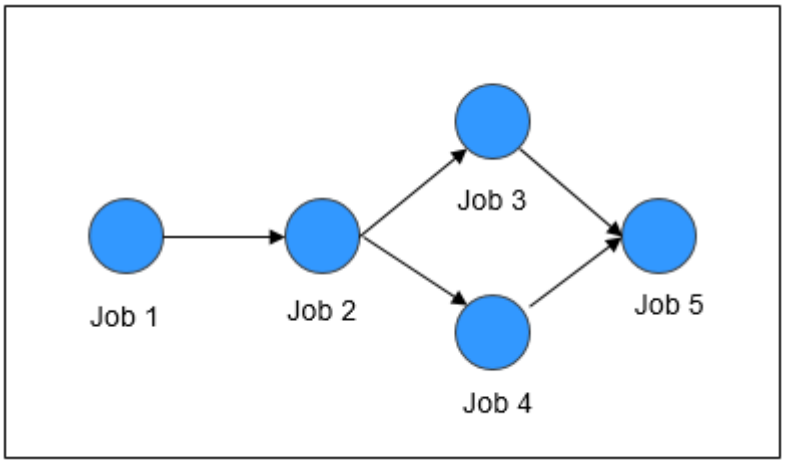
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Operators are connected using the TCP protocol. In distributed environment, operators can be scheduled to any node and upstream and downstream services cannot detect the scheduling.

Job Pipeline

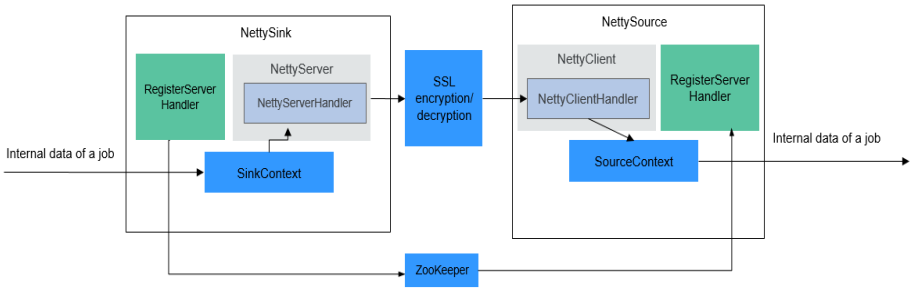
A pipeline consists of multiple Flink jobs connected through TCP. Upstream jobs can send data to downstream jobs. The flow diagram about data transmission is called a job pipeline, as shown in **Figure 1-24**.

Figure 1-24 Job pipeline



Job Pipeline Principles

Figure 1-25 Job Pipeline



- NettySink and NettySource

  In a pipeline, upstream jobs and downstream jobs communicate with each other through Netty. The Sink operator of the upstream job works as a server and the Source operator of the downstream job works as a client. The Sink operator of the upstream job is called NettySink, and the Source operator of the downstream job is called NettySource.

- NettyServer and NettyClient

  NettySink functions as the server of Netty. In NettySink, NettyServer achieves the function of a server. NettySource functions as the client of Netty. In NettySource, NettyClient achieves the function of a client.

- Publisher

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The job that sends data to downstream jobs through NettySink is called a publisher.

- Subscriber

  The job that receives data from upstream jobs through NettySource is called a subscriber.

- RegisterServer

  RegisterServer is the third-party memory that stores the IP address, port number, and concurrency information about NettyServer.

- The general outside-in architecture is as follows:
  - NettySink->NettyServer->NettyServerHandler
  - NettySource->NettyClient->NettyClientHandler

Job Pipeline Functions

- **NettySink**

  NettySink consists of the following major modules:

  - RichParallelSinkFunction

    NettySink inherits RichParallelSinkFunction and attributes of Sink operators. The RichParallelSinkFunction API implements following functions:

    - Starts the NettySink operator.
    - Runs the NettySink operator and receives data from the upstream operator.
    - Cancels the running of NettySink operators.

    The following information can be obtained using the attribute of RichParallelSinkFunction:

    - subtaskIndex about the concurrency of each NettySink operator
    - Concurrency of the NettySink operator

  - RegisterServerHandler

    RegisterServerHandler interacts with the component of RegisterServer and defines following APIs:

    - **start();**: Starts the RegisterServerHandler and establishes a contact with the third-party RegisterServer.
    - **createTopicNode();**: Creates a topic node.
    - **register();**: Registers information such as the IP address, port number, and concurrency to the topic node.
    - **deleteTopicNode();**: Deletes a topic node.
    - **unregister();**: Deletes registration information.
    - **query();**: Queries registration information.
    - **isExist();**: Verifies that a specific piece of information exists.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **shutdown();**: Disables the RegisterServerHandler and disconnects from the third-party RegisterServer.

  NOTE

  - RegisterServerHandler API enables ZooKeeper to work as the handler of RegisterServer. You can customize your handler as required. Information is stored in ZooKeeper in the following form:

    ```
    Namespace
    |---Topic-1
      |---parallel-1
      |---parallel-2
      |....
      |---parallel-n
    |---Topic-2
      |---parallel-1
      |---parallel-2
      |....
      |---parallel-m
    |...
    ```

  - Information about NameSpace can be obtained from the following parameters of the **flink-conf.yaml** file:

    ```
    nettyconnector.registerserver.topic.storage: /flink/nettyconnector
    ```

  - The simple authentication and security layer (SASL) authentication between ZookeeperRegisterServerHandler and ZooKeeper is implemented through the Flink framework.
  - Ensure that each job has a unique topic. Otherwise, the subscription relationship may be unclear.
  - When calling **shutdown()**, ZookeeperRegisterServerHandler deletes the registration information about the current concurrency, and then attempts to delete the topic node. If the topic node is not empty, deletion will be canceled, because not all concurrency has exited.

– NettyServer

  NettyServer is the core of the NettySink operator, whose main function is to create a NettyServer and receive connection requests from NettyClient. Use NettyServerHandler to send data received from upstream operators of a same job. The port number and subnet of NettyServer needs to be configured in the **flink-conf.yaml** file.

  - Port range

    ```
    nettyconnector.sinkserver.port.range: 28444-28843
    ```

  - Subnet

    ```
    nettyconnector.sinkserver.subnet: 10.162.222.123/24
    ```

    NOTE

    The **nettyconnector.sinkserver.subnet** parameter is set to the subnet (service IP address) of the Flink client by default. If the client and TaskManager are not in the same subnet, an error may occur. Therefore, you need to manually set this parameter to the subnet (service IP address) of TaskManager.

– NettyServerHandler

  The handler enables the interaction between NettySink and subscribers. After NettySink receives messages, the handler sends these messages out. To ensure data transmission security, this channel is encrypted using SSL. The **nettyconnector.ssl.enabled** configures whether to enable SSL encryption. The SSL encryption is enabled only when **nettyconnector.ssl.enabled** is set to **true**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **NettySource**

  NettySource consists of the following major modules:

  - RichParallelSourceFunction

    NettySource inherits RichParallelSinkFunction and attributes of Source operators. The RichParallelSourceFunction API implements following functions:

    - Starts the NettySink operator.
    - Runs the NettySink operator, receives data from subscribers, and injects the data to jobs.
    - Cancels the running of Source operators.

    The following information can be obtained using the attribute of RichParallelSinkFunction:

    - subtaskIndex about the concurrency of each NettySource operator
    - Concurrency of the NettySource operator

    When the NettySource operator enters the running stage, the NettyClient status is monitored. Once abnormality occurs, NettyClient is restarted and reconnected to NettyServer, preventing data confusion.

  - RegisterServerHandler

    RegisterServerHandler of NettySource has similar functions as the RegisterServerHandler of NettySink. It obtains the IP address, port number, and information of concurrent operators of each subscribed job obtained in the NettySource operator.

  - NettyClient

    NettyClient establishes a connection with NettyServer and uses NettyClientHandler to receive data. Each NettySource operator must have a unique name (specified by the user). NettyServer determines whether each client comes from different NettySources based on unique names. When a connection is established between NettyClient and NettyServer, NettyClient is registered with NettyServer and the NettySource name of NettyClient is transferred to NettyServer.

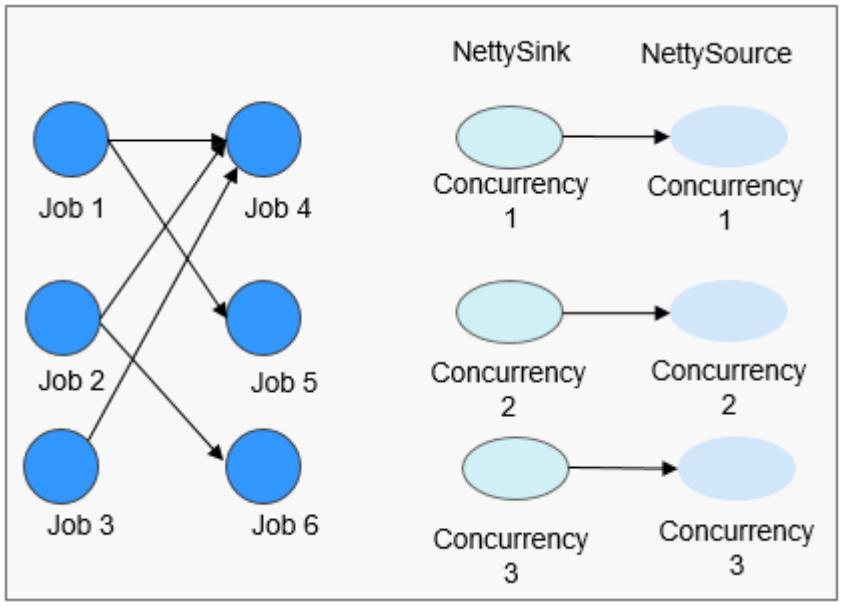
  - NettyClientHandler

    The NettyClientHandler enables the interaction with publishers and other operators of the job. When messages are received, NettyClientHandler transfers these messages to the job. To ensure secure data transmission, SSL encryption is enabled for the communication with NettySink. The SSL encryption is enabled only when SSL is enabled and **nettyconnector.ssl.enabled** is set to **true**.

The relationship between the jobs may be many-to-many. The concurrency between each NettySink and NettySource operator is one-to-many, as shown in **Figure 1-26**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-26 Relationship diagram



1.6.6.6 Flink Stream SQL Join Enhancement

Flink's Table API&SQL is an integrated query API for Scala and Java that allows the composition of queries from relational operators such as selection, filter, and join in an intuitive way.

Introduction to Stream SQL Join

SQL Join is used to query data based on the relationship between columns in two or more tables. Flink Stream SQL Join allows you to join two streaming tables and query results from them. Queries similar to the following are supported:

```
SELECT o.proctime, o.productId, o.orderId, s.proctime AS shipTime
FROM Orders AS o
JOIN Shipments AS s
ON o.orderId = s.orderId
AND o.proctime BETWEEN s.proctime AND s.proctime + INTERVAL '1' HOUR;
```

Currently, Stream SQL Join needs to be performed within a specified window. The join operation for data within the window requires at least one equi-join predicate and a join condition that bounds the time on both sides. Such a condition can be defined by two appropriate range predicates (**<**, **<=**, **>=**, **>**), a **BETWEEN** predicate, or a single equality predicate that compares the same type of time attributes (such as processing time or event time) of both input tables.

The following example will join all orders with their corresponding shipments if the order was shipped four hours after the order was received.

```
SELECT *
FROM Orders o, Shipments s
WHERE o.id = s.orderId AND
o.ordertime BETWEEN s.shiptime - INTERVAL '4' HOUR AND s.shiptime
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

1. Stream SQL Join supports only inner join.
2. The **ON** clause should include an equal join condition.
3. Time attributes support only the processing time and event time.
4. The window condition supports only the bounded time range, for example, **o.proctime BETWEEN s.proctime - INTERVAL '1' HOUR AND s.proctime + INTERVAL '1' HOUR**. The unbounded range such as **o. proctime > s.proctime** is not supported. The **proctime** attribute of two streams must be included. **o.proctime BETWEEN proctime () AND proctime () + 1** is not supported.

1.6.6.7 Flink CEP in SQL Enhancement

Flink CEP in SQL

Flink allows users to represent complex event processing (CEP) query results in SQL for pattern matching and evaluate event streams on Flink engines.

SQL Query Syntax

CEP SQL is implemented through the **MATCH_RECOGNIZE** SQL syntax. The **MATCH_RECOGNIZE** clause is supported by Oracle SQL since Oracle Database 12c and is used to indicate event pattern matching in SQL. Apache Calcite also supports the **MATCH_RECOGNIZE** clause.

Flink uses Calcite to analyze SQL query results. Therefore, this operation complies with the Apache Calcite syntax.

```
MATCH_RECOGNIZE (
    [ PARTITION BY expression [, expression ]* ]
    [ ORDER BY orderItem [, orderItem ]* ]
    [ MEASURES measureColumn [, measureColumn ]* ]
    [ ONE ROW PER MATCH | ALL ROWS PER MATCH ]
    [ AFTER MATCH
          ( SKIP TO NEXT ROW
          | SKIP PAST LAST ROW
          | SKIP TO FIRST variable
          | SKIP TO LAST variable
          | SKIP TO variable )
    ]
    PATTERN ( pattern )
    [ WITHIN intervalLiteral ]
    [ SUBSET subsetItem [, subsetItem ]* ]
    DEFINE variable AS condition [, variable AS condition ]*
    )
```

The syntax elements of the **MATCH_RECOGNIZE** clause are defined as follows:

(Optional) **-PARTITION BY**: defines partition columns. This clause is optional. If this parameter is not defined, the parallelism 1 is used.

(Optional) **-ORDER BY**: defines the sequence of events in a data flow. The **ORDER BY** clause is optional. If it is ignored, non-deterministic sorting is used. Since the order of events is important in pattern matching, this clause should be specified in most cases.

(Optional) **-MEASURES**: specifies the attribute value of the successfully matched event.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

(Optional) **-ONE ROW PER MATCH | ALL ROWS PER MATCH**: defines how to output the result. **ONE ROW PER MATCH** indicates that only one row is output for each matching. **ALL ROWS PER MATCH** indicates that one row is output for each matching event.

(Optional) **-AFTER MATCH**: specifies the start position for processing after the next pattern is successfully matched.

-PATTERN: defines the matching pattern as a regular expression. The following operators can be used in the **PATTERN** clause: join operators, quantifier operators (*, +, ?, {n}, {n,}, {n,m}, and {,m}), branch operators (vertical bar |), and differential operators ('{- -}').

(Optional) **-WITHIN**: outputs a pattern clause match only when the match occurs within the specified time.

(Optional) **-SUBSET**: combines one or more associated variables defined in the **DEFINE** clause.

-DEFINE: specifies the Boolean condition, which defines the variables used in the **PATTERN** clause.

In addition, the **MATCH_RECOGNIZE** clause supports the following functions:

-MATCH_NUMBER(): Used in the **MEASURES** clause to allocate the same number to each row that is successfully matched.

-CLASSIFIER(): Used in the **MEASURES** clause to indicate the mapping between matched rows and variables.

-FIRST() and **LAST()**: Used in the **MEASURES** clause to return the value of the expression evaluated in the first or last row of the row set mapped to the schema variable.

-NEXT() and **PREV()**: Used in the **DEFINE** clause to evaluate an expression using the previous or next row in a partition.

-RUNNING and **FINAL** keywords: Used to determine the semantics required for aggregation. **RUNNING** can be used in the **MEASURES** and **DEFINE** clauses, whereas **FINAL** can be used only in the **MEASURES** clause.

- Aggregate functions (**COUNT**, **SUM**, **AVG**, **MAX**, **MIN**): Used in the **MEASURES** and **DEFINE** clauses.

Query Example

The following query finds the V-shaped pattern in the stock price data flow.

```
SELECT *
   FROM MyTable
   MATCH_RECOGNIZE (
     ORDER BY rowtime
     MEASURES
       STRT.name as s_name,
       LAST(DOWN.name) as down_name,
       LAST(UP.name) as up_name
     ONE ROW PER MATCH
     PATTERN (STRT DOWN+ UP+)
     DEFINE
       DOWN AS DOWN.v < PREV(DOWN.v),
       UP AS UP.v > PREV(UP.v)
   )
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In the following query, the aggregate function **AVG** is used in the **MEASURES** clause of **SUBSET E** consisting of variables related to A and C.

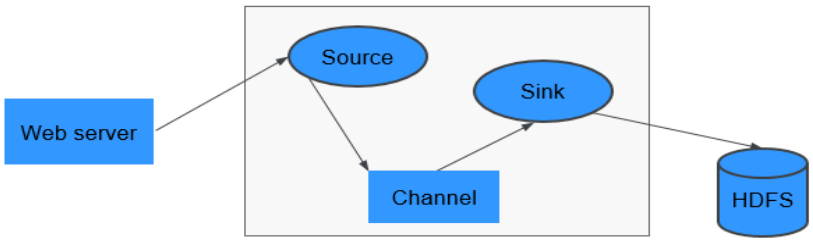
```
SELECT *
   FROM Ticker
   MATCH_RECOGNIZE (
     MEASURES
       AVG(E.price) AS avgPrice
     ONE ROW PER MATCH
     AFTER MATCH SKIP PAST LAST ROW
     PATTERN (A B+ C)
     SUBSET E = (A,C)
     DEFINE
       A AS A.price < 30,
       B AS B.price < 20,
       C AS C.price < 30
   )
```

1.6.7 Flume

1.6.7.1 Flume Basic Principles

Flume is a distributed, reliable, and HA system that supports massive log collection, aggregation, and transmission. Flume supports customization of various data senders in the log system for data collection. In addition, Flume can roughly process data and write data to various data receivers (customizable). A Flume-NG is a branch of Flume. It is simple, small, and easy to deploy. The following figure shows the basic architecture of the Flume-NG.

Figure 1-27 Flume-NG architecture



A Flume-NG consists of agents. Each agent consists of three components (source, channel, and sink). A source is used for receiving data. A channel is used for transmitting data. A sink is used for sending data to the next end.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

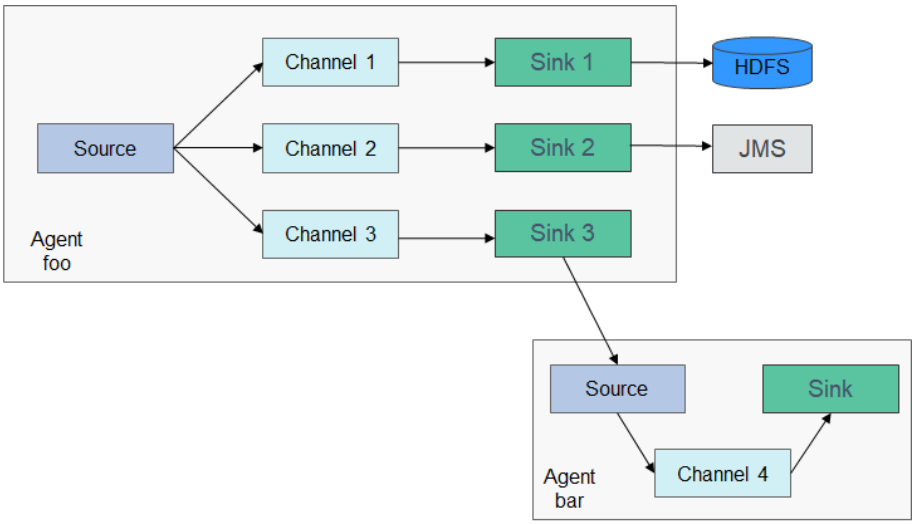
Table 1-6 Module description

Module	Description
Source	A source receives data or generates data by using a special mechanism, and places the data in batches in one or more channels. The source can work in data-driven or polling mode. Typical source types are as follows: • Sources that are integrated with the system, such as Syslog and Netcat • Sources that automatically generate events, such as Exec and SEQ • IPC sources that are used for communications between agents, such as Avro A source must be associated with at least one channel.
Channel	A channel is used to buffer data between a source and a sink. The channel caches data from the source and deletes that data after the sink sends the data to the next-hop channel or final destination. Different channels provide different persistence levels. • Memory channel: non-persistency • File channel: Write-Ahead Logging (WAL) based persistence • JDBC channel: persistency implemented based on the embedded database The channel supports the transaction feature to ensure simple sequential operations. A channel can work with sources and sinks of any quantity.
Sink	A sink sends data to the next-hop channel or final destination. Once completed, the transmitted data is removed from the channel. Typical sink types are as follows: • Sinks that send storage data to the final destination, such as HDFS and HBase • Sinks that are consumed automatically, such as Null Sink • IPC Sinks used for communication between Agents, such as Avro A sink must be associated with a specific channel.

As shown in **Figure 1-28**, a Flume client can have multiple sources, channels, and sinks.

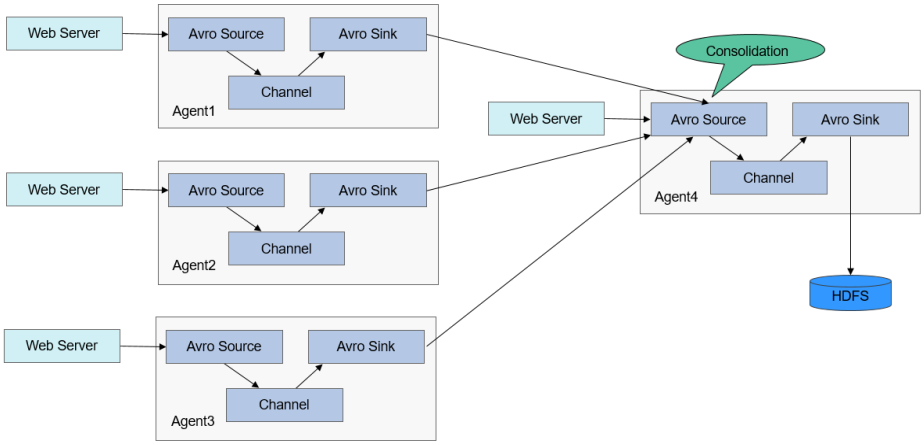
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-28 Flume structure



The reliability of Flume depends on transaction switchovers between agents. If the next agent is abnormal, the channel stores data persistently and transmits data until the agent recovers. The availability of Flume depends on the built-in load balancing and failover mechanisms. Both the channel and agent can be configured with multiple entities between which they can use load balancing policies. Each agent is a Java Virtual Machine (JVM) process. A server can have multiple agents. Collection nodes (for example, Agents 1, 2, 3) process logs. Aggregation nodes (for example, Agent 4) write the logs into HDFS. The agent of each collection node can select multiple aggregation nodes for load balancing.

Figure 1-29 Flume cascading



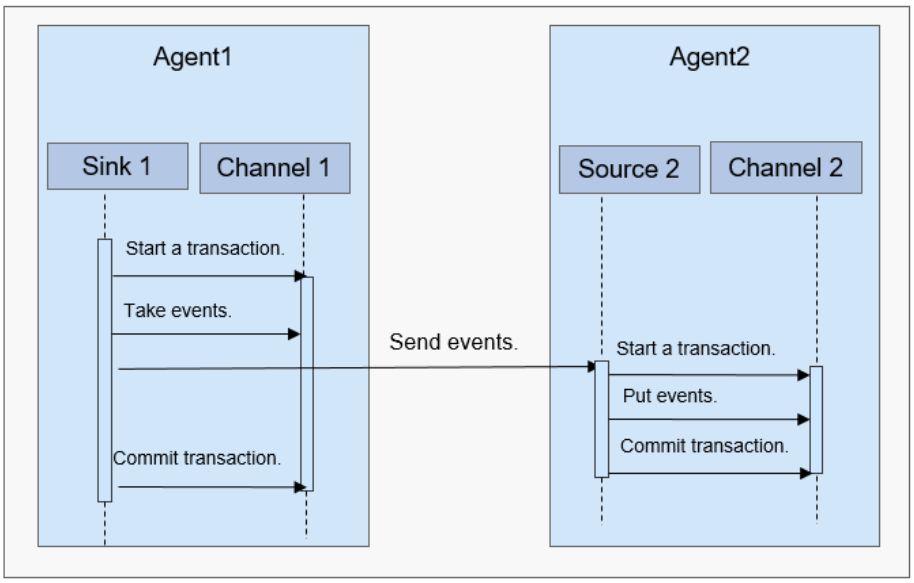
Principle

Reliability Between Agents

Figure 1-30 shows the data exchange between agents.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-30 Data transmission process



1. Flume ensures reliable data transmission based on transactions. When data flows from one agent to another agent, the two transactions take effect. The sink of Agent 1 (agent that sends a message) needs to obtain a message from a channel and sends the message to Agent 2 (agent that receives the message). If Agent 2 receives and successfully processes the message, Agent 1 will submit a transaction, indicating a successful and reliable data transmission.
2. When receiving a message from Agent 1, Agent 2 starts a new transaction. If the message is processed successfully (written to a channel), Agent 2 submits the transaction and sends a success response to Agent 1.
3. If the data transmission fails before a commit operation, the last transaction starts again and retransmits the data that fails to be transmitted last time. The commit operation has written the transaction into a disk. Therefore, the last transaction can continue after the process fails and restores.

1.6.7.2 Relationships Between Flume and Other Components

Relationship Between Flume and HDFS

If HDFS is configured as the Flume sink, HDFS functions as the final data storage system of Flume. Flume installs, configures, and writes all transmitted data into HDFS.

Relationship Between Flume and HBase

If HBase is configured as the Flume sink, HBase functions as the final data storage system of Flume. Flume writes all transmitted data into HBase based on configurations.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.7.3 Flume Enhanced Open Source Features

Flume Enhanced Open Source Features

- Improving transmission speed: Multiple lines instead of only one line of data can be specified as an event. This improves the efficiency of code execution and reduces the times of disk writes.
- Transferring ultra-large binary files: According to the current memory usage, Flume automatically adjusts the memory used for transferring ultra-large binary files to prevent out-of-memory.
- Supporting the customization of preparations before and after transmission: Flume supports customized scripts to be run before or after transmission for making preparations.
- Managing client alarms: Flume receives Flume client alarms through MonitorServer and reports the alarms to the alarm management center on MRS Manager.

1.6.8 Guardian

Guardian Basic Principles

Guardian is a service that provides temporary authentication credentials for services such as HDFS, Hive, Spark, Loader and HetuEngine to access OBS in decoupled storage and compute scenarios. The Guardian component needs to be installed only when OBS is connected. Typical features of Guardian include:

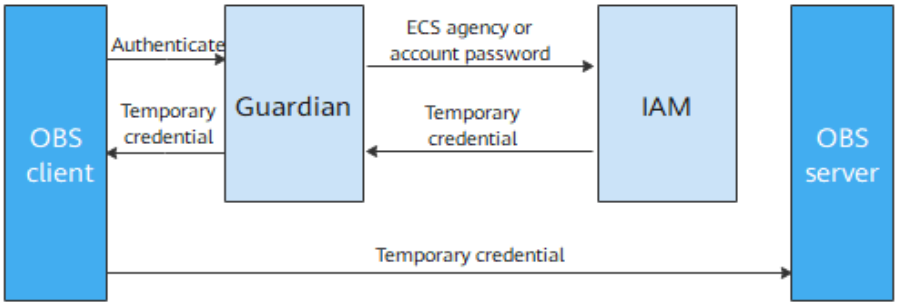
- Provides the capability of obtaining temporary authentication credentials for accessing OBS.
- Provides fine-grained permission control for accessing OBS.
- Provides the unified cache refreshing capability for temporary authentication credentials used to access OBS.

The Guardian server provides functions for the TokenServer role. TokenServer supports multi-instance deployment. Each instance can have the same functions. A single point of failure (SPOF) does not affect service functions. In addition, the Guardian server provides RPC and HTTPS interfaces to obtain temporary authentication credentials for accessing OBS.

Guardian Architecture

Figure 1-31 shows the basic architecture of Guardian.

Figure 1-31 Guardian architecture

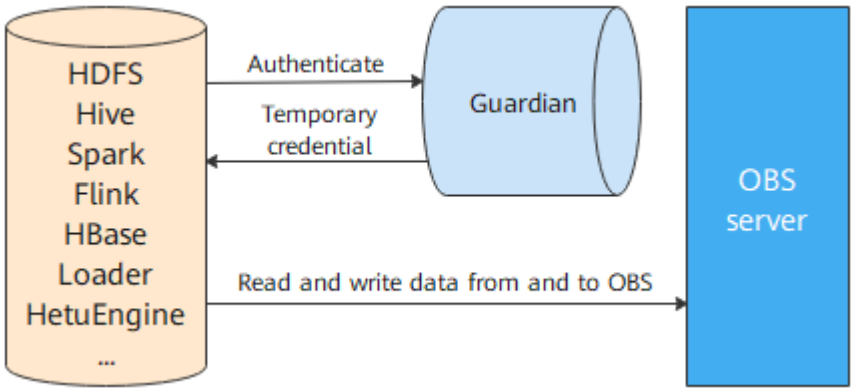


 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Relationships Between Guardian and Other Components

Before accessing OBS, HDFS, Hive, Spark, Flink, Loader, and HetuEngine access Guardian to obtain temporary credentials for the access. Guardian generates a temporary credential with fine-grained authentication content based on the IAM access request of the current login user and returns the credential to the component. The component uses the credential to access OBS. OBS determines whether the current user has the access permission based on the credential.

Figure 1-32 Relationships between Guardian and other components



1.6.9 HBase

1.6.9.1 HBase Basic Principles

HBase undertakes data storage. HBase is an open source, column-oriented, distributed storage system that is suitable for storing massive amounts of unstructured or semi-structured data. It features high reliability, high performance, and flexible scalability, and supports real-time data read/write. For more information about HBase, see **https://hbase.apache.org/**.

Typical features of a table stored in HBase are as follows:

- Big table (BigTable): One table contains hundred millions of rows and millions of columns.
- Column-oriented: Column-oriented storage, retrieval, and permission control
- Sparse: Null columns in the table do not occupy any storage space.

MRS HBase supports secondary indexing to allow indexes to be created for column values so that data can be filtered by column using native HBase APIs.

HBase Architecture

An HBase cluster consists of active and standby Master processes and multiple RegionServer processes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-33 HBase architecture

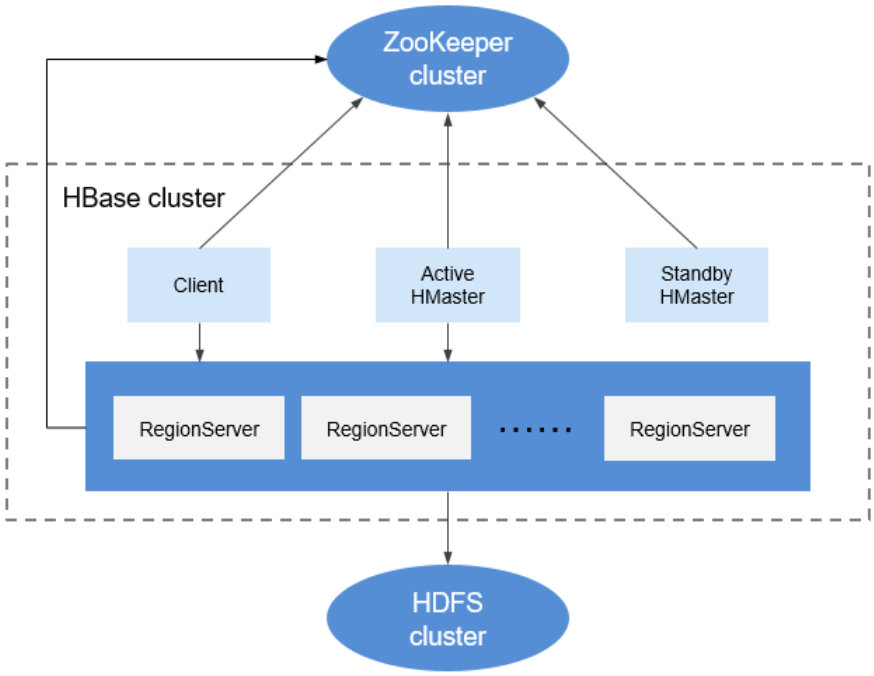


Table 1-7 Module description

Module	Description
Master	Master is also called HMaster. In HA mode, HMaster consists of an active HMaster and a standby HMaster. • Active Master: manages RegionServer in HBase, including the creation, deletion, modification, and query of a table, balances the load of RegionServer, adjusts the distribution of Region, splits Region and distributes Region after it is split, and migrates Region after RegionServer expires. • Standby Master: takes over services when the active HMaster is faulty. The original active HMaster demotes to the standby HMaster after the fault is rectified.
Client	Client communicates with Master for management and with RegionServer for data protection by using the Remote Procedure Call (RPC) mechanism of HBase.
RegionServer	RegionServer provides read and write services of table data as a data processing and computing unit in HBase. RegionServer is deployed with DataNodes of HDFS clusters to store data.
ZooKeeper cluster	ZooKeeper provides distributed coordination services for processes in HBase clusters. Each RegionServer is registered with ZooKeeper so that the active Master can obtain the health status of each RegionServer.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
HDFS cluster	HDFS provides highly reliable file storage services for HBase. All HBase data is stored in the HDFS.

HBase Principles

- **HBase Data Model**

  HBase stores data in tables, as shown in **Figure 1-34**. Data in a table is divided into multiple Regions, which are allocated by Master to RegionServers for management.

  Each Region contains data within a RowKey range. An HBase data table contains only one Region at first. As the number of data increases and reaches the upper limit of the Region capacity, the Region is split into two Regions. You can define the RowKey range of a Region when creating a table or define the Region size in the configuration file.

  Figure 1-34 HBase data model

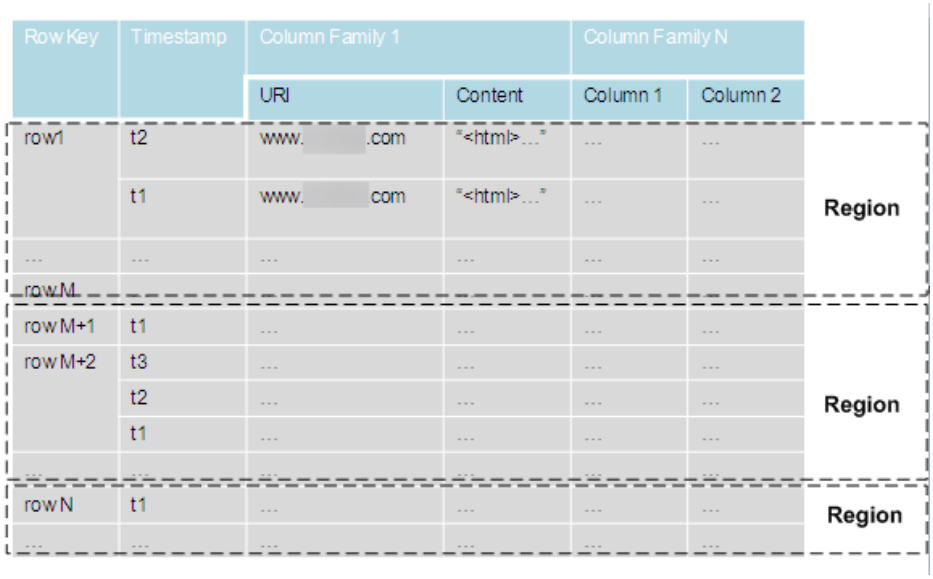


Table 1-8 Concepts

Module	Description
RowKey	Similar to the primary key in a relationship table, which is the unique ID of the data in each row. A RowKey can be a string, integer, or binary string. All records are stored after being sorted by RowKey.
Timestamp	The timestamp of a data operation. Data can be specified with different versions by time stamp. Data of different versions in each cell is stored by time in descending order.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
Cell	Minimum storage unit of HBase, consisting of keys and values. A key consists of six fields, namely row, column family, column qualifier, timestamp, type, and MVCC version. Values are the binary data objects.
Column Family	One or multiple horizontal column families form a table. A column family can consist of multiple random columns. A column is a label under a column family, which can be added as required when data is written. The column family supports dynamic expansion so the number and type of columns do not need to be predefined. Columns of a table in HBase are sparsely distributed. The number and type of columns in different rows can be different. Each column family has the independent time to live (TTL). You can lock the row only. Operations on the row in a column family are the same as those on other rows.
Column	Similar to traditional databases, HBase tables also use columns to store data of the same type.

- **RegionServer Data Storage**

  RegionServer manages the regions allocated by HMaster. **Figure 1-35** shows the data storage structure of RegionServer.

  Figure 1-35 RegionServer data storage structure

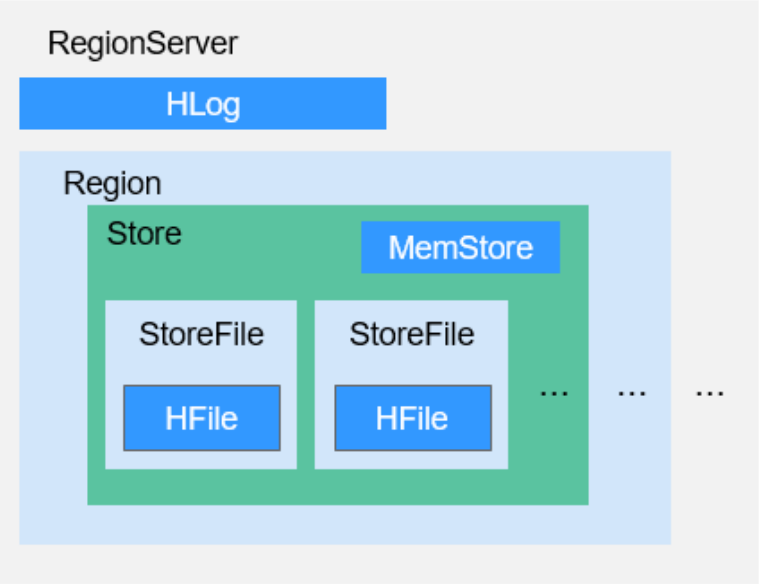


  Table 1-9 lists each component of Region described in **Figure 1-35**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-9 Region structure description

Module	Description
Store	A Region consists of one or multiple Stores. Each Store maps a column family in **Figure 1-34**.
MemStore	A Store contains one MemStore. The MemStore caches data inserted to a Region by the client. When the MemStore capacity reaches the upper limit, RegionServer flushes data in MemStore to the HDFS.
StoreFile	The data flushed to the HDFS is stored as a StoreFile in the HDFS. As more data is inserted, multiple StoreFiles are generated in a Store. When the number of StoreFiles reaches the upper limit, RegionServer merges multiple StoreFiles into a big StoreFile.
HFile	HFile defines the storage format of StoreFiles in a file system. HFile is the underlying implementation of StoreFile.
HLog	HLogs prevent data loss when RegionServer is faulty. Multiple Regions in a RegionServer share the same HLog.

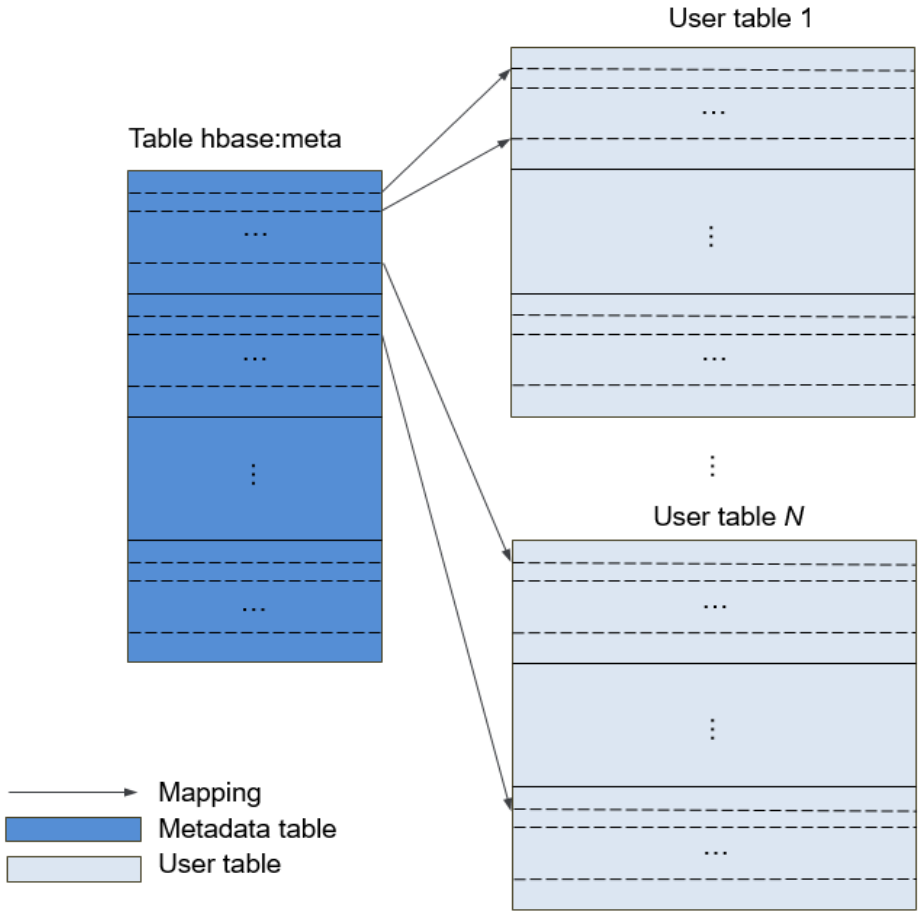
- **Metadata Table**

  The metadata table is a special HBase table, which is used by the client to locate a region. Metadata table includes **hbase:meta** table to record region information of user tables, such as the region location and start and end RowKey.

  Figure 1-36 shows the mapping relationship between metadata tables and user tables.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

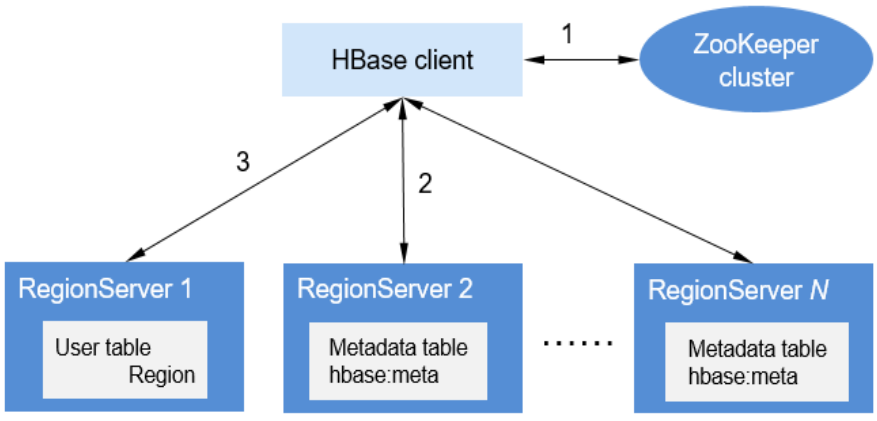
Figure 1-36 Mapping relationships between metadata tables and user tables



- **Data Operation Process**

  Figure 1-37 shows the HBase data operation process.

  Figure 1-37 Data processing



  a. When you add, delete, modify, and query HBase data, the HBase client first connects to ZooKeeper to obtain information about the RegionServer where the **hbase:meta** table is located. If you modify the NameSpace, such as creating and deleting a table, you need to access HMaster to update the meta information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

b. The HBase client connects to the RegionServer where the region of the **hbase:meta** table is located and obtains the RegionServer location where the region of the user table resides.

c. Then the HBase client connects to the RegionServer where the region of the user table is located and issues a data operation command to the RegionServer. The RegionServer executes the command.

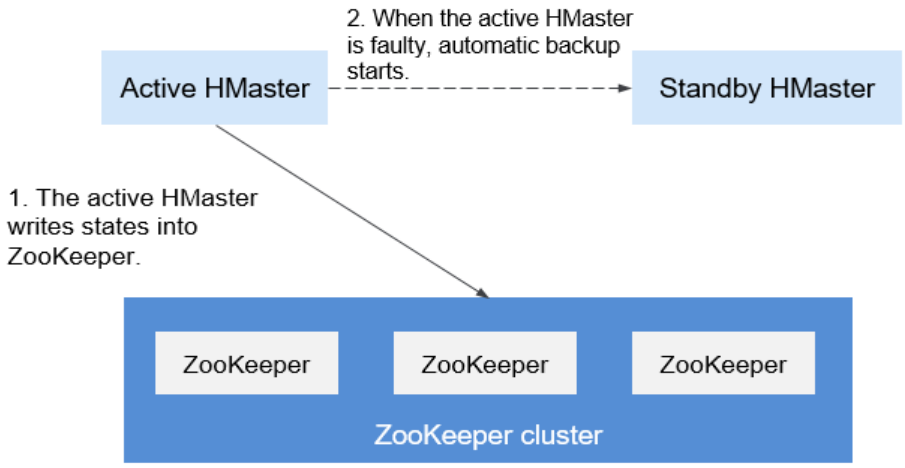
To improve data processing efficiency, the HBase client caches region information of the **hbase:meta** table and user table. When an application initiates a second data operation, the HBase client queries the region information from the memory. If no match is found in the memory, the HBase client performs the preceding operations to obtain region information.

1.6.9.2 HBase HA Solution

HBase HA

HMaster in HBase allocates Regions. When one RegionServer service is stopped, HMaster migrates the corresponding Region to another RegionServer. The HMaster HA feature is brought in to prevent HBase functions from being affected by the HMaster single point of failure (SPOF).

Figure 1-38 HMaster HA implementation architecture



The HMaster HA architecture is built based on Ephemeral nodes (temporary nodes) created in the ZooKeeper cluster.

Upon startup, HMaster nodes try to create a master znode in the ZooKeeper cluster. The HMaster node that creates the master znode first becomes the active HMaster, and the other is the standby HMaster.

It will add watch events to the master node. If the service on the active HMaster is stopped, the active HMaster disconnects from the ZooKeeper cluster. After the session expires, the active HMaster disappears. The standby HMaster detects the disappearance of the active HMaster through watch events and creates a master node to make itself be the active one. Then, the active/standby switchover

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

completes. If the failed node detects existence of the master node after being restarted, it enters the standby state and adds watch events to the master node.

When the client accesses the HBase, it first obtains the HMaster's address based on the master node information on the ZooKeeper and then establishes a connection to the active HMaster.

1.6.9.3 Relationship Between HBase and Other Components

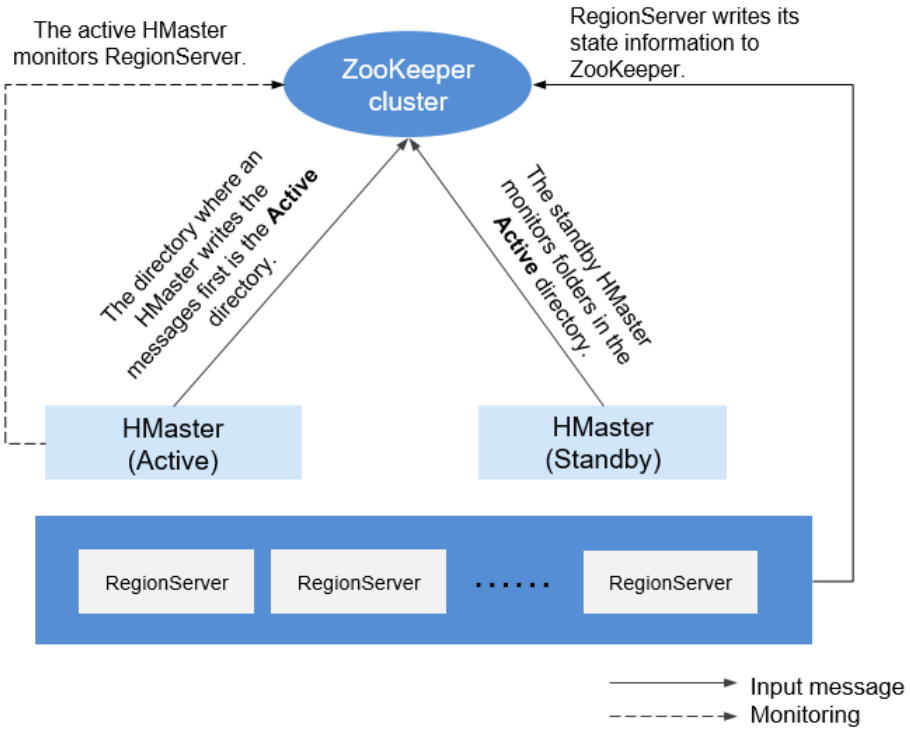
Relationship Between HDFS and HBase

HDFS is the subproject of Apache Hadoop. HBase uses the Hadoop Distributed File System (HDFS) as the file storage system. HBase is located in structured storage layer. The HDFS provides highly reliable support for lower-layer storage of HBase. All the data files of HBase can be stored in the HDFS, except some log files generated by HBase.

Relationship Between ZooKeeper and HBase

Figure 1-39 describes the relationship between ZooKeeper and HBase.

Figure 1-39 Relationship between ZooKeeper and HBase



1. HRegionServer registers itself to ZooKeeper in Ephemeral node. ZooKeeper stores the HBase information, including the HBase metadata and HMaster addresses.
2. HMaster detects the health status of each HRegionServer using ZooKeeper, and monitors them.
3. HBase supports multiple HMaster nodes (like HDFS NameNodes). When the active HMatser is faulty, the standby HMaster obtains the state information

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

about the entire cluster using ZooKeeper. This prevents a failure of the HBase cluster.

1.6.9.4 HBase Enhanced Open Source Features

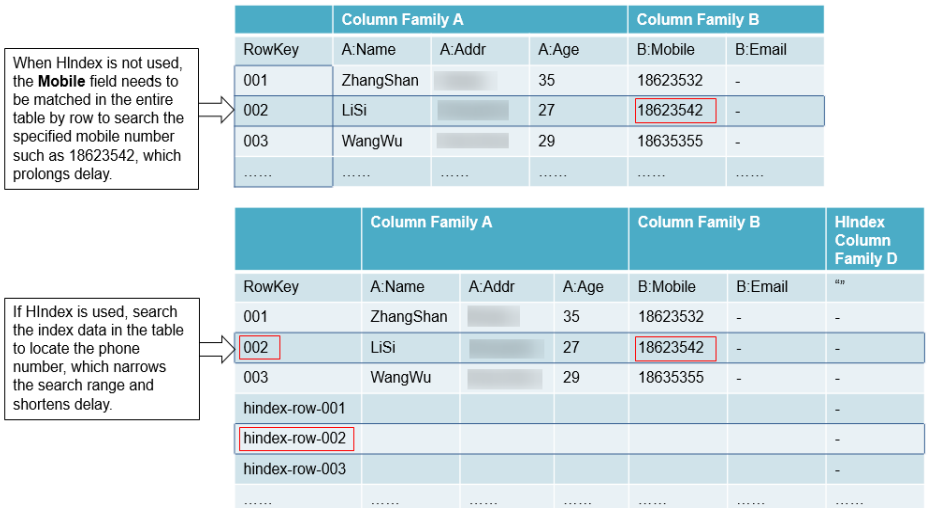
HIndex

HBase is a distributed Key-Value store. Data of a table is sorted in the alphabetic order based on row keys. If you query data based on a specified row key or scan data in the scale of a specified row key, HBase can quickly locate the target data, enhancing the efficiency.

However, in most actual scenarios, you need to query the data of which the column value is *XXX*. HBase provides the Filter feature to query data with a specific column value. All data is scanned in the order of row keys, and then the data is matched with the specific column value until the required data is found. The Filter feature scans some unnecessary data to obtain the only required data. Therefore, the Filter feature cannot meet the requirements of frequent queries with high performance standards.

HBase HIndex is designed to address these issues. HBase HIndex enables HBase to query data based on specific column values.

Figure 1-40 HIndex



- Rolling upgrade is not supported for index data.
- Restrictions of combined indexes:
  - All columns involved in combined indexes must be entered or deleted in a single mutation. Otherwise, inconsistency will occur.

    Index: **IDX1=>cf1:[q1->datatype],[q2];cf2:[q2->datatype]**

    Correct write operations:

    ```
    Put put = new Put(Bytes.toBytes("row"));
    put.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q1"), Bytes.toBytes("valueA"));
    put.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q2"), Bytes.toBytes("valueB"));
    put.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q2"), Bytes.toBytes("valueC"));
    table.put(put);
    ```

    Incorrect write operations:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
Put put1 = new Put(Bytes.toBytes("row"));
put1.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q1"), Bytes.toBytes("valueA"));
table.put(put1);
Put put2 = new Put(Bytes.toBytes("row"));
put2.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q2"), Bytes.toBytes("valueB"));
table.put(put2);
Put put3 = new Put(Bytes.toBytes("row"));
put3.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q2"), Bytes.toBytes("valueC"));
table.put(put3);
```

- The combined conditions-based query is supported only when the combined index column contains filter criteria, or StartRow and StopRow are not specified for some index columns.

  Index: **IDX1=>cf1:[q1->datatype],[q2];cf2:[q1->datatype]**

  Correct query operations:

  ```
  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',>=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf1','q2',>=,'binary:valueB',true,true) AND
  SingleColumnValueFilter('cf2','q1',>=,'binary:valueC',true,true) "}

  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf1','q2',>=,'binary:valueB',true,true)" }

  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',>=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf1','q2',>=,'binary:valueB',true,true) AND
  SingleColumnValueFilter('cf2','q1',>=,'binary:valueC',true,true)",STARTROW=>'row001',STOPROW
  =>'row100'}
  ```

  Incorrect query operations:

  ```
  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',>=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf1','q2',>=,'binary:valueB',true,true) AND
  SingleColumnValueFilter('cf2','q1',>=,'binary:valueC',true,true)  AND
  SingleColumnValueFilter('cf2','q2',>=,'binary:valueD',true,true)"}

  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf2','q1',>=,'binary:valueC',true,true)" }

  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf2','q2',>=,'binary:valueD',true,true)" }

  scan 'table', {FILTER=>"SingleColumnValueFilter('cf1','q1',=,'binary:valueA',true,true) AND
  SingleColumnValueFilter('cf1','q2',>=,'binary:valueB',true,true)" ,STARTROW=>'row001',STOPROW
  =>'row100' }
  ```

- Do not configure any split policy for tables with index data.
- Other mutation operations, such as **increment** and **append**, are not supported.
- Index of the column with **maxVersions** greater than 1 is not supported.
- The data index column in a row cannot be updated.

  Index 1: **IDX1=>cf1:[q1->datatype],[q2];cf2:[q1->datatype]**

  Index 2: **IDX2=>cf2:[q2->datatype]**

  Correct update operations:

  ```
  Put put1 = new Put(Bytes.toBytes("row"));
  put1.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q1"), Bytes.toBytes("valueA"));
  put1.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q2"), Bytes.toBytes("valueB"));
  put1.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q1"), Bytes.toBytes("valueC"));
  put1.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q2"), Bytes.toBytes("valueD"));
  table.put(put1);

  Put put2 = new Put(Bytes.toBytes("row"));
  put2.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q3"), Bytes.toBytes("valueE"));
  put2.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q3"), Bytes.toBytes("valueF"));
  table.put(put2);
  ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Incorrect update operations:

```
Put put1 = new Put(Bytes.toBytes("row"));
put1.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q1"), Bytes.toBytes("valueA"));
put1.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q2"), Bytes.toBytes("valueB"));
put1.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q1"), Bytes.toBytes("valueC"));
put1.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q2"), Bytes.toBytes("valueD"));
table.put(put1);

Put put2 = new Put(Bytes.toBytes("row"));
put2.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q1"), Bytes.toBytes("valueA_new"));
put2.addColumn(Bytes.toBytes("cf1"), Bytes.toBytes("q2"), Bytes.toBytes("valueB_new"));
put2.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q1"), Bytes.toBytes("valueC_new"));
put2.addColumn(Bytes.toBytes("cf2"), Bytes.toBytes("q2"), Bytes.toBytes("valueD_new"));
table.put(put2);
```

- The table to which an index is added cannot contain a value greater than 32 KB.
- If user data is deleted due to the expiration of the column-level TTL, the corresponding index data is not deleted immediately. It will be deleted in the major compaction operation.
- The TTL of the user column family cannot be modified after the index is created.
  - If the TTL of a column family increases after an index is created, delete the index and re-create one. Otherwise, some generated index data will be deleted before user data is deleted.
  - If the TTL value of the column family decreases after an index is created, the index data will be deleted after user data is deleted.
- The index query does not support the reverse operation, and the query results are disordered.
- The index does not support the **clone snapshot** operation.
- Index tables must use HIndexWALPlayer to replay logs. WALPlayer cannot be used to replay logs.

```
hbase org.apache.hadoop.hbase.hindex.mapreduce.HIndexWALPlayer
Usage: WALPlayer [options] <wal inputdir> <tables> [<tableMappings>]
Read all WAL entries for <tables>.
If no tables ("") are specific, all tables are imported.
(Careful, even -ROOT- and hbase:meta entries will be imported in that case.)
Otherwise <tables> is a comma separated list of tables.

The WAL entries can be mapped to new set of tables via <tableMapping>.
<tableMapping> is a command separated list of targettables.
If specified, each table in <tables> must have a mapping.

By default WALPlayer will load data directly into HBase.
To generate HFiles for a bulk data load instead, pass the option:
  -Dwal.bulk.output=/path/for/output
  (Only one table can be specified, and no mapping is allowed!)
Other options: (specify time range to WAL edit to consider)
  -Dwal.start.time=[date|ms]
  -Dwal.end.time=[date|ms]
For performance also consider the following options:
  -Dmapreduce.map.speculative=false
  -Dmapreduce.reduce.speculative=false
```

- When the **deleteall** command is executed for the index table, the performance is low.
- The index table does not support HBCK. To use HBCK to repair an index table, delete index data first.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

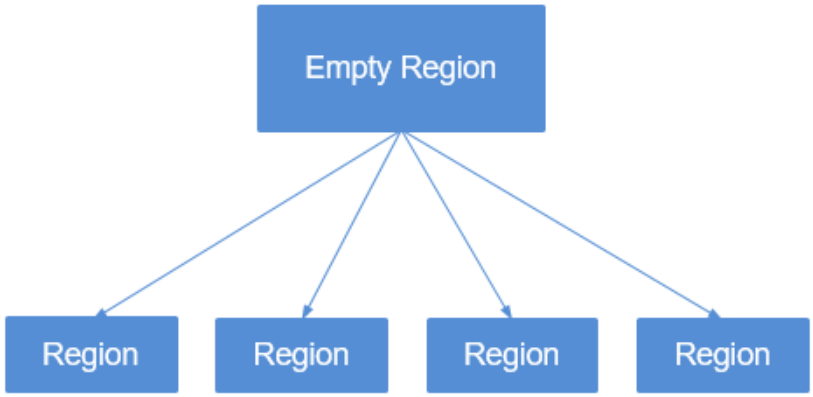
Multi-point Division

When you create tables that are pre-divided by region in HBase, you may not know the data distribution trend so the division by region may be inappropriate. After the system runs for a period, regions need to be divided again to achieve better performance. Only empty regions can be divided.

The region division function delivered with HBase divides regions only when they reach the threshold. This is called "single point division".

To achieve better performance when regions are divided based on user requirements, multi-point division is developed, which is also called "dynamic division". That is, an empty region is pre-divided into multiple regions to prevent performance deterioration caused by insufficient region space.

Figure 1-41 Multi-point division



Connection Limitation

Too many sessions mean that too many queries and MapReduce tasks are running on HBase, which compromises HBase performance and even causes service rejection. You can configure parameters to limit the maximum number of sessions that can be established between the client and the HBase server to achieve HBase overload protection.

Improved Disaster Recovery

The disaster recovery (DR) capabilities between the active and standby clusters can enhance HA of the HBase data. The active cluster provides data services and the standby cluster backs up data. If the active cluster is faulty, the standby cluster takes over data services. Compared with the open source replication function, this function is enhanced as follows:

1. The whitelist function is enabled for the standby cluster. Data can be pushed to clusters only by specifying IP addresses.
2. In the open source version, replication is synchronized based on WAL, and data backup is implemented by replaying WAL in the standby cluster. For BulkLoad operations, since no WAL is generated, data will not be replicated to the standby cluster. By recording BulkLoad operations on the WAL and

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

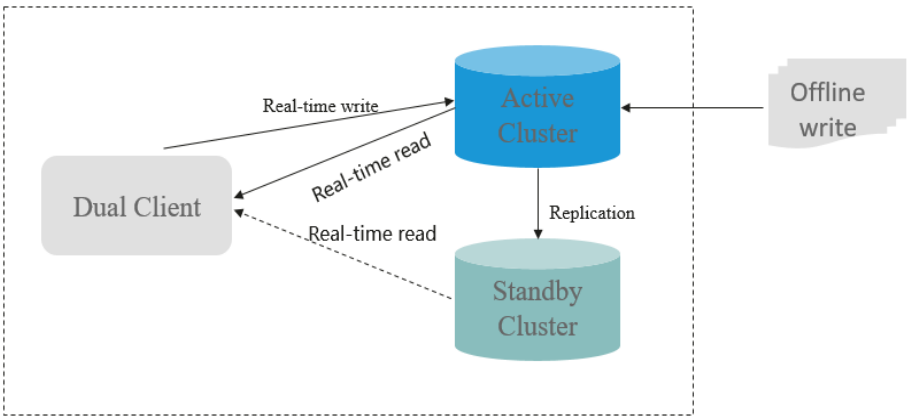
synchronizing them to the standby cluster, the standby cluster can read BulkLoad operation records through WAL and load HFile in the active cluster to the standby cluster to implement data backup.

3. In the open source version, HBase filters ACLs. Therefore, ACL information will not be synchronized to the standby cluster. By adding a filter (**org.apache.hadoop.hbase.replication.SystemTableWALEntryFilterAllowACL**), ACL information can be synchronized to the standby cluster. You can configure **hbase.replication.filter.sytemWALEntryFilter** to enable the filter and implement ACL synchronization.
4. The standby cluster is read-only for HBase clients that are not deployed on the standby cluster. Only the internal administrative user of the nodes in the standby cluster can modify HBase.

HBase Dual-Read

In the HBase storage scenario, it is difficult to ensure 99.9% query stability due to GC, network jitter, and bad sectors of disks. The HBase dual-read feature is added to meet the requirements of low glitches during large-data-volume random read.

The HBase dual-read feature is based on the DR capability of the active and standby clusters. The probability that the two clusters generate glitches at the same time is far less than that of one cluster. The dual-cluster concurrent access mode is used to ensure query stability. When a user initiates a query request, the HBase service of the two clusters is queried at the same time. If the active cluster does not return any result after a period of time (the maximum tolerable glitch time), the data of the cluster with the fastest response can be used. The following figure shows the working principle.



Custom Delimiters Supported on Phoenix CsvBulkLoadTool

NOTE

This feature is available in MRS 3.2.0 or later.

Currently, Phoenix's open source CsvBulkLoadTool supports only a single character as the data delimiter. When a user data file contains any characters, a special string is used as the delimiter. To meet this requirement, custom delimiters are supported so you can use any visible characters within the specified length as delimiters to import data files.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

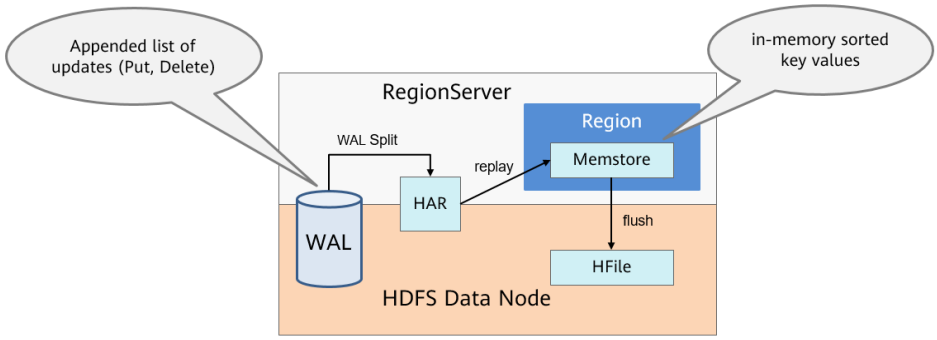
Writing Small Files Generated During WAL File Splitting to the HTTP Archive (HAR) File

NOTE

This feature is available in MRS 3.2.0 or later.

When a RegionServer is faulty or restarted, HMaster uses ServerCrashProcedure to restore the services running on the RegionServer. The restoration process involves splitting WAL files. During WAL file splitting, a large number of small files are generated, which may cause HDFS performance bottlenecks. As a result, service restoration takes a long time.

This feature writes small files to the HAR file during WAL file splitting to shorten the RegionServer restoration duration.



Batch TRSP

HBase 2.*x* uses HBase Procedure to rewrite the region assignment logic (AMV2). When each region is opened or closed, a TransitRegionStateProcedure (TRSP) is associated with it. When services running on a RegionServer need to be restored due to RegionServer faults or restarts, HMaster creates a TRSP for each region to be restored. A large number of TRSPs need to persist data to Proc WAL files and perform an RPC interaction with RegionServer, which may cause HMaster performance bottlenecks. As a result, the service restoration takes a long time.

This feature attaches regions to TRSPs and uses one TRSP to restore all regions of a RegionServer. RegionServer batch opens or closes regions and reports all regions to HMaster at a time.

NOTE

This feature can only restore regions to their original RegionServers. Therefore, the prerequisite for this optimization to take effect is that the faulty or restarted RegionServer has been brought online again when HMaster creates a TRSP. This feature is used to optimize the duration for HBase restart or service fault restoration. If a few RegionServers are faulty, this feature may not take effect because HMaster had created TRSPs before RegionServers were brought online again.

This feature is available in MRS 3.2.0 or later.

Self-Healing from HBase Hotspotting

This function applies to MRS 3.3.0 and later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

HBase is a distributed key-value database. Regions are the smallest units for HBase data management. Poorly designed row keys make requests directed at a few fixed regions. As a result, the service pressure is high on a node, causing performance deterioration or even request failures.

MetricController instances are added to HBase. After hotspotting detection is enabled, the request traffic of each RegionServer node is monitored. Through aggregation analysis, the nodes and regions with excessive requests can be identified, accelerating hotspotting detection. In addition, this self-healing function transfers workload or performs region splitting. For hotspotting caused by a single rowkey and sequential writes where the self-healing function does not work, traffic limiting is provided instead to minimize the impact on other services on a node.

In MRS 3.5.0 and later versions, the MetricController instance is optional. If MetricController is required, it must be deployed in active/standby mode and share the same node with HMaster. If the MetricController instance is deployed, hotspotting detection is enabled by default and the request traffic of each RegionServer node can be monitored. Through aggregation analysis, the nodes and regions with excessive requests can be identified, helping quickly identify hotspotting. In addition, the self-healing from hotspotting function is provided to transfer workload or perform region splitting. For hotspotting caused by a single rowkey and sequential writes where the self-healing function does not work, traffic limiting is provided instead to minimize the impact on other services on a node.

In MRS 3.5.0 and later versions, the following self-healing events and alarms may be generated:

- "Hotspot Region Splitting" event: Hotspot regions are split. The names of the processed tables and regions are displayed in the additional information of this event.
- "Hotspot Region Isolation" event: By limiting the requests to the hotspot region, the RPC requests received by the hotspot region are isolated to an extra thread pool. The number of handlers used for processing the requests is specified by **hbase.regionserver.hotregion.handler.count**. The names of the processed tables and regions are displayed in the additional information of this event.
- "Hotspot Region Migrated to other ResionServers" event: The hotspot region is moved to other RegionServers to handle the overwhelmed requests. The names of the processed tables and regions are displayed in the additional information of this event.
- "Region Traffic Restriction for HBase" alarm: This alarm is triggered by a hotspot region isolation event. This alarm is generated when the system detects that there are region requests in the thread pool of hotspot traffic restriction on RegionServer. The additional information contains the name of the table and region for which the alarm is generated. For details, see the handling procedure of "ALM-19023 Region Traffic Restriction for HBase" in the alarm reference.

To view the events or alarms, log in to FusionInsight Manager and choose **O&M** > **Alarms** > **Events** or **Alarms**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- To disable automatic recovery for HBase hotspotting, log in to FusionInsight Manager, click **Cluster**, choose **Services** > **HBase**, and click **Configurations**. Search for **hbase.metric.report.provider** and **hbase.rowkey.metric.provider**, leave the two parameters empty, and click **Save**. Click **Instances**, select all MetricController instances, and choose **More** > **Stop Instance**.
- For MRS 3.5.0-LTS or later, to enable the HBase hotspot self-healing function, log in to FusionInsight Manager, choose **Cluster** > **Services** > **HBase** > **Configurations**, search for the following parameters, and change their values:
  - Set **hbase.metric.report.provider** to the default value **org.apache.hadoop.hbase.regionserver.hotspot.HotspotMetricCollectorImpl**.
  - Set **hbase.rowkey.metric.provider** to the default value **org.apache.hadoop.hbase.regionserver.hotspot.RowkeyMetricsImpl**.

  Click **Save**. Click **Instances**, select all MetricController instances, and click **Start Instance**.

1.6.10 HDFS

1.6.10.1 HDFS Basic Principles

Hadoop Distributed File System (HDFS) implements reliable and distributed read/write of massive amounts of data. HDFS is applicable to the scenario where data read/write features "write once and read multiple times". However, the write operation is performed in sequence, that is, it is a write operation performed during file creation or an adding operation performed behind the existing file. HDFS ensures that only one caller can perform write operation on a file but multiple callers can perform read operation on the file at the same time.

NOTE

To use HDFS, ensure that the Hadoop service has been installed in the MRS cluster.

Architecture

HDFS consists of active and standby NameNodes and multiple DataNodes, as shown in **Figure 1-42**.

HDFS works in master/slave architecture. NameNodes run on the master (active) node, and DataNodes run on the slave (standby) node. ZKFC should run along with the NameNodes.

The communication between NameNodes and DataNodes is based on Transmission Control Protocol (TCP)/Internet Protocol (IP). The NameNode, DataNode, ZKFC, and JournalNode can be deployed on Linux servers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-42 HA HDFS architecture

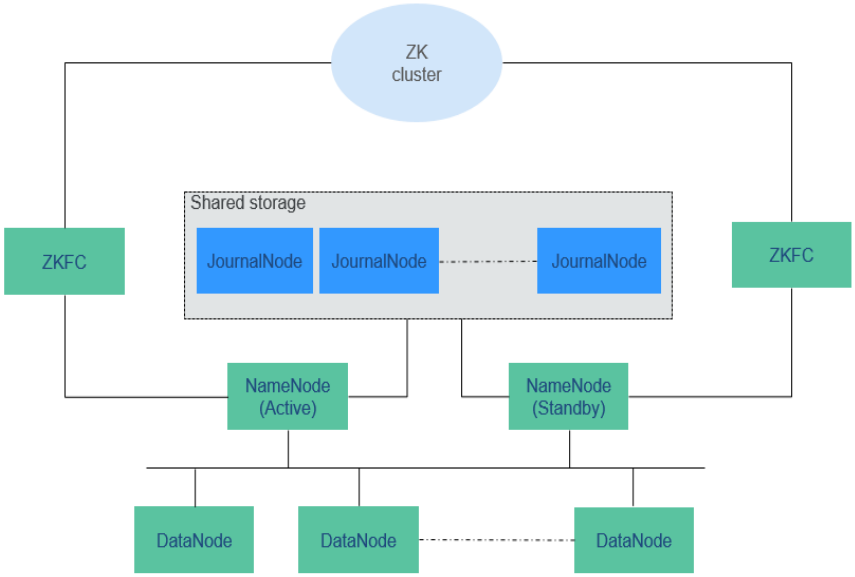


Table 1-10 describes the functions of each module shown in **Figure 1-42**.

Table 1-10 Module description

Module	Description
NameNode	A NameNode is used to manage the namespace, directory structure, and metadata information of a file system and provide the backup mechanism. The NameNode is classified into the following two types: • Active NameNode: manages the namespace, maintains the directory structure and metadata of file systems, and records the mapping relationships between data blocks and files to which the data blocks belong. • Standby NameNode: synchronizes with the data in the active NameNode, and takes over services from the active NameNode when the active NameNode is faulty. • Observer NameNode: synchronizes with the data in the active NameNode, and processes read requests from the client.
DataNode	A DataNode is used to store data blocks of each file and periodically report the storage status to the NameNode.
JournalNode	In HA cluster, synchronizes metadata between the active and standby NameNodes.
ZKFC	ZKFC must be deployed for each NameNode. It monitors NameNode status and writes status information to ZooKeeper. ZKFC also has permissions to select the active NameNode.
ZK Cluster	ZooKeeper is a coordination service which helps the ZKFC to elect the active NameNode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
HttpFS gateway	HttpFS is a single stateless gateway process which provides the WebHDFS REST API for external processes and FileSystem API for the HDFS. HttpFS is used for data transmission between different versions of Hadoop. It is also used as a gateway to access the HDFS behind a firewall.

- **HDFS HA Architecture**

  HA is used to resolve the SPOF problem of NameNode. This feature provides a standby NameNode for the active NameNode. When the active NameNode is faulty, the standby NameNode can quickly take over to continuously provide services for external systems.

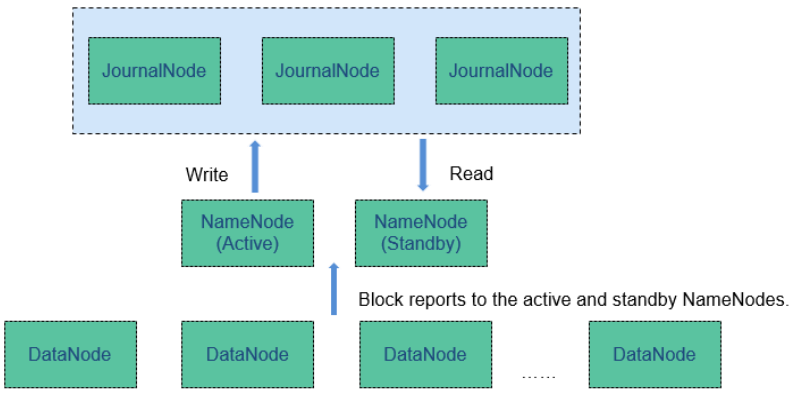
  In a typical HDFS HA scenario, there are usually two NameNodes. One is in the active state, and the other in the standby state.

  A shared storage system is required to support metadata synchronization of the active and standby NameNodes. This version provides Quorum Journal Manager (QJM) HA solution, as shown in **Figure 1-43**. A group of JournalNodes are used to synchronize metadata between the active and standby NameNodes.

  Generally, an odd number (2N+1) of JournalNodes are configured, and at least three JournalNodes are required. For one metadata update message, data writing is considered successful as long as data writing is successful on N +1 JournalNodes. In this case, data writing failure of a maximum of N JournalNodes is allowed. For example, when there are three JournalNodes, data writing failure of one JournalNode is allowed; when there are five JournalNodes, data writing failure of two JournalNodes is allowed.

  JournalNode is a lightweight daemon process and shares a host with other services of Hadoop. It is recommended that the JournalNode be deployed on the control node to prevent data writing failure on the JournalNode during massive data transmission.

  Figure 1-43 QJM-based HDFS architecture



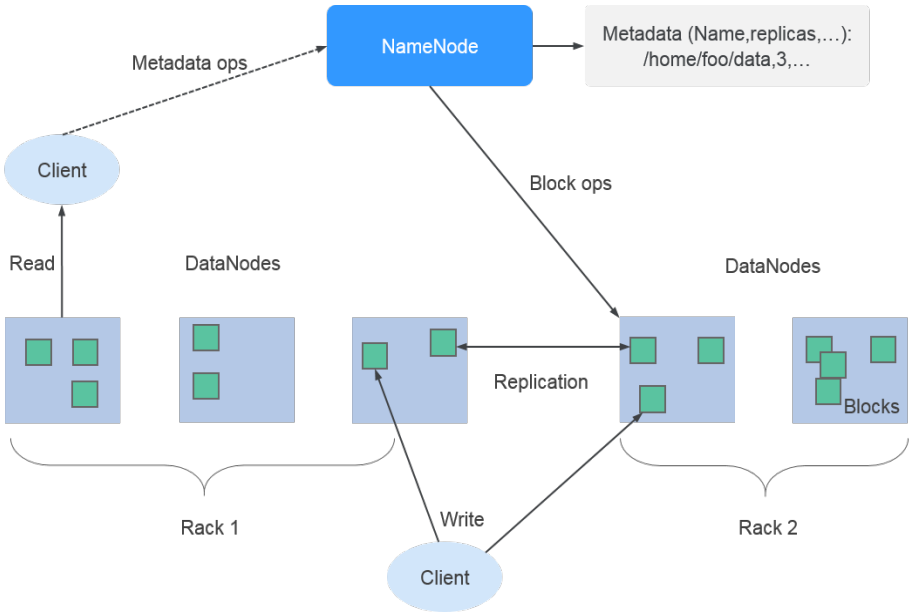
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

HDFS Reliability

MRS uses the HDFS copy mechanism to ensure data reliability. One backup file is automatically generated for each file saved in HDFS, that is, two copies are generated in total. The number of HDFS copies can be queried using the **dfs.replication** parameter.

- When the Core node specification of the MRS cluster is set to non-local hard disk drive (HDD) and the cluster has only one Core node, the default number of HDFS copies is 1. If the number of Core nodes in the cluster is greater than or equal to 2, the default number of HDFS copies is 2.
- When the Core node specification of the MRS cluster is set to local disk and the cluster has only one Core node, the default number of HDFS copies is 1. If there are two Core nodes in the cluster, the default number of HDFS copies is 2. If the number of Core nodes in the cluster is greater than or equal to 3, the default number of HDFS copies is 3.

Figure 1-44 HDFS architecture



Supports scheduling for node balance on HDFS and disk balance on a single node, improving HDFS storage performance after node or disk scale-out.

1.6.10.2 HDFS HA Solution

HDFS HA Background

In versions earlier than Hadoop 2.0.0, SPOF occurs in the HDFS cluster. Each cluster has only one NameNode. If the host where the NameNode is located is faulty, the HDFS cluster cannot be used unless the NameNode is restarted or started on another host. This affects the overall availability of HDFS in the following aspects:

1. In the case of an unplanned event such as host breakdown, the cluster would be unavailable until the NameNode is restarted.

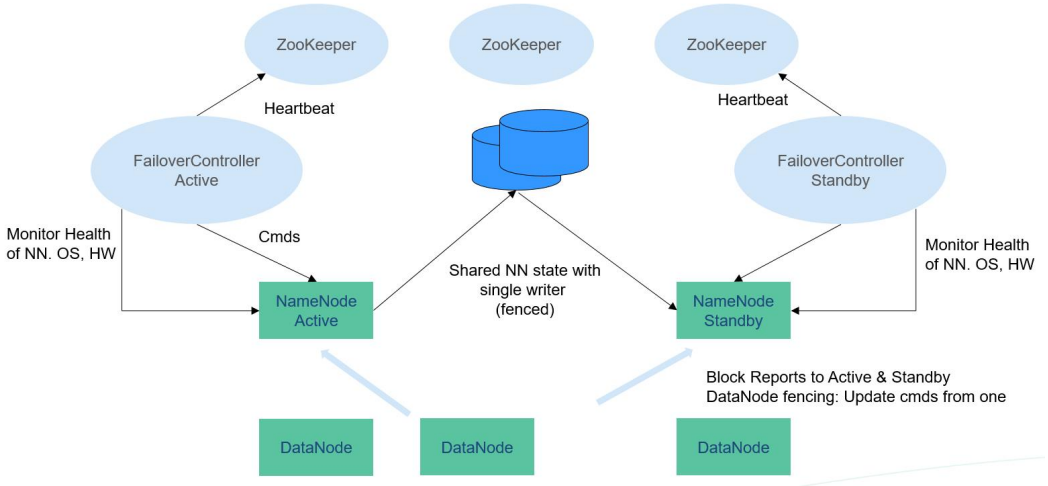
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Planned maintenance tasks, such as software and hardware upgrade, will cause the cluster stop working.

To solve the preceding problems, the HDFS HA solution enables a hot-swap NameNode backup for NameNodes in a cluster in automatic or manual (configurable) mode. When a machine fails (due to hardware failure), the active/standby NameNode switches over automatically in a short time. When the active NameNode needs to be maintained, the MRS cluster administrator can manually perform an active/standby NameNode switchover to ensure cluster availability during maintenance.

HDFS HA Implementation

Figure 1-45 Typical HA deployment



In a typical HA cluster (as shown in **Figure 1-45**), two NameNodes need to be configured on two independent servers, respectively. At any time point, one NameNode is in the active state, and the other NameNode is in the standby state. The active NameNode is responsible for all client operations in the cluster, while the standby NameNode maintains synchronization with the active node to provide fast switchover if necessary.

To keep the data synchronized with each other, both nodes communicate with a group of JournalNodes. When the active node modifies any file system's metadata, it will store the modification log to a majority of these JournalNodes. For example, if there are three JournalNodes, then the log will be saved on two of them at least. The standby node monitors changes of JournalNodes and synchronizes changes from the active node. Based on the modification log, the standby node applies the changes to the metadata of the local file system. Once a switchover occurs, the standby node can ensure its status is the same as that of the active node. This ensures that the metadata of the file system is synchronized between the active and standby nodes if the switchover is incurred by the failure of the active node.

To ensure fast switchover, the standby node needs to have the latest block information. Therefore, DataNodes send block information and heartbeat messages to two NameNodes at the same time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

It is vital for an HA cluster that only one of the NameNodes be active at any time. Otherwise, the namespace state would split into two parts, risking data loss or other incorrect results. To prevent the so-called "split-brain scenario", the JournalNodes will only ever allow a single NameNode to write data to it at a time. During switchover, the NameNode which is to become active will take over the role of writing data to JournalNodes. This effectively prevents the other NameNodes from being in the active state, allowing the new active node to safely proceed with switchover.

1.6.10.3 Relationship Between HDFS and Other Components

Relationship Between HDFS and HBase

HDFS is a subproject of Apache Hadoop, which is used as the file storage system for HBase. HBase is located in the structured storage layer. HDFS provides highly reliable support for lower-layer storage of HBase. All the data files of HBase can be stored in the HDFS, except some log files generated by HBase.

Relationship Between HDFS and MapReduce

- HDFS features high fault tolerance and high throughput, and can be deployed on low-cost hardware for storing data of applications with massive data sets.
- MapReduce is a programming model used for parallel computation of large data sets (larger than 1 TB). Data computed by MapReduce comes from multiple data sources, such as Local FileSystem, HDFS, and databases. Most data comes from the HDFS. The high throughput of HDFS can be used to read massive data. After being computed, data can be stored in HDFS.

Relationship Between HDFS and Spark

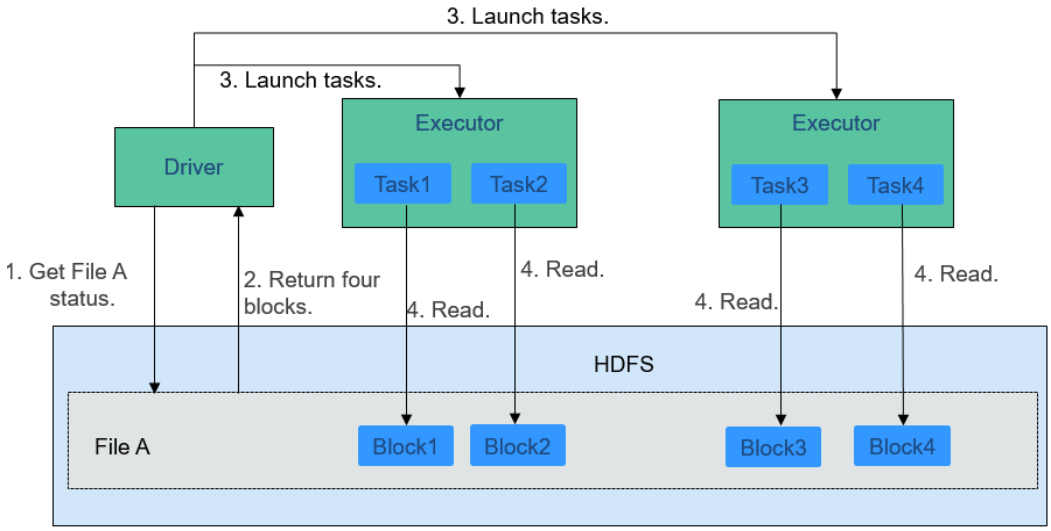
Data computed by Spark comes from multiple data sources, such as local files and HDFS. Most data comes from HDFS which can read data in large scale for parallel computing. After being computed, data can be stored in HDFS.

Spark involves Driver and Executor. Driver schedules tasks and Executor runs tasks.

Figure 1-46 shows how data is read from a file.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-46 File reading process

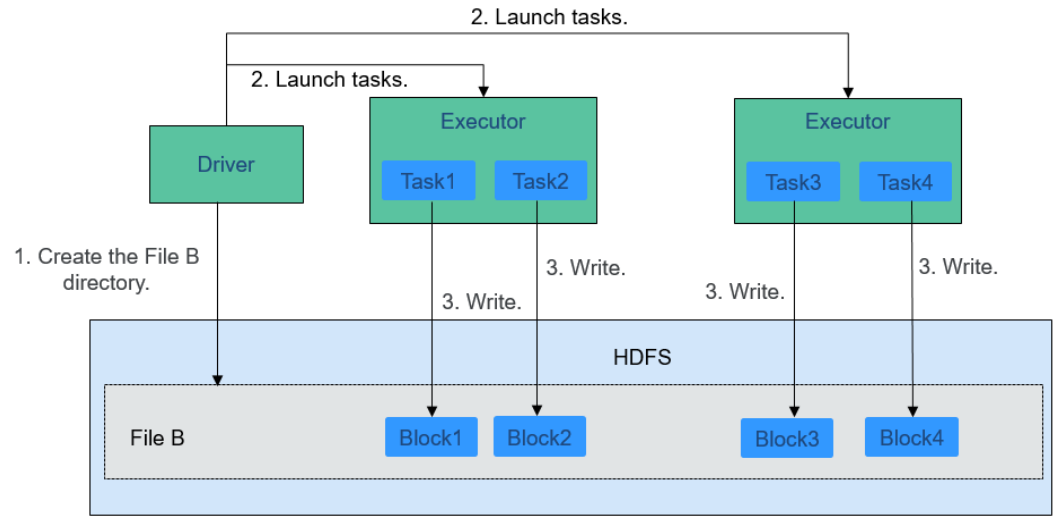


The file reading process is as follows:

1. Driver interconnects with HDFS to obtain the information of File A.
2. The HDFS returns the detailed block information about this file.
3. Driver sets a parallel degree based on the block data amount, and creates multiple tasks to read the blocks of this file.
4. Executor runs the tasks and reads the detailed blocks as part of the Resilient Distributed Dataset (RDD).

Figure 1-47 shows how data is written to a file.

Figure 1-47 File writing process



The file writing process is as follows:

1. Driver creates a directory where the file is to be written.
2. Based on the RDD distribution status, the number of tasks related to data writing is computed, and these tasks are sent to Executor.

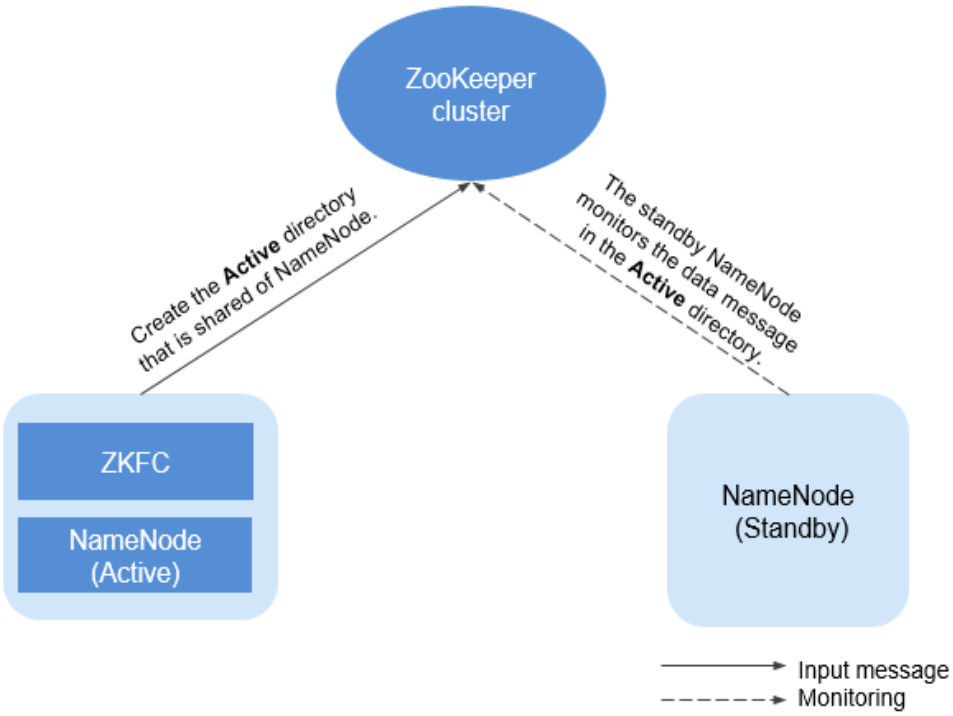
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Executor runs these tasks, and writes the computed RDD data to the directory created in 1.

Relationship Between HDFS and ZooKeeper

Figure 1-48 shows the relationship between ZooKeeper and HDFS.

Figure 1-48 Relationship between ZooKeeper and HDFS



As the client of a ZooKeeper cluster, ZKFailoverController (ZKFC) monitors the status of NameNode. ZKFC is deployed only in the node where NameNode resides, and in both the active and standby HDFS NameNodes.

1. The ZKFC connects to ZooKeeper and saves information such as host names to ZooKeeper under the znode directory **/hadoop-ha**. NameNode that creates the directory first is considered as the active node, and the other is the standby node. NameNodes read the NameNode information periodically through ZooKeeper.
2. When the process of the active node ends abnormally, the standby NameNode detects changes in the **/hadoop-ha** directory through ZooKeeper, and then takes over the service of the active NameNode.

1.6.10.4 HDFS Enhanced Open Source Features

Enhanced Open Source Feature: File Block Colocation

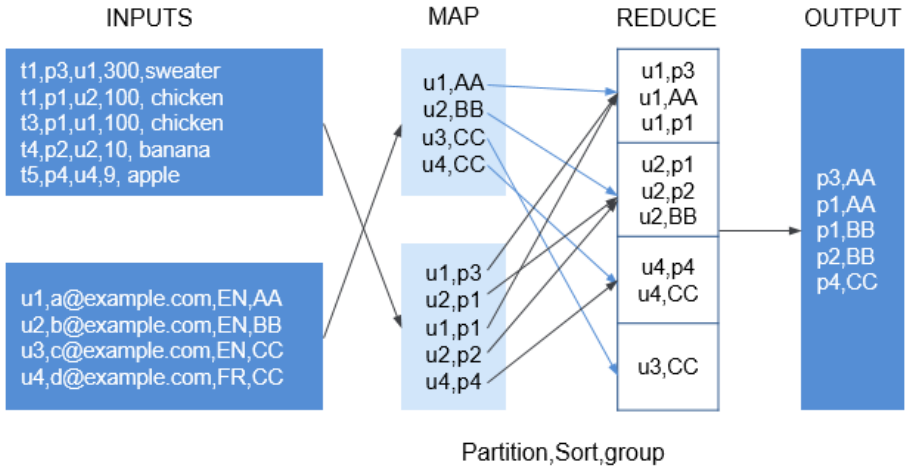
In the offline data summary and statistics scenario, Join is a frequently used computing function, and is implemented in MapReduce as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. The Map task processes the records in the two table files into Join Key and Value, performs hash partitioning by Join Key, and sends the data to different Reduce tasks for processing.
2. Reduce tasks read data in the left table recursively in the nested loop mode and traverse each line of the right table. If join key values are identical, join results are output.

   The preceding method sharply reduces the performance of the join calculation. Because a large amount of network data transfer is required when the data stored in different nodes is sent from MAP to Reduce, as shown in **Figure 1-49**.

Figure 1-49 Data transmission in the non-colocation scenario

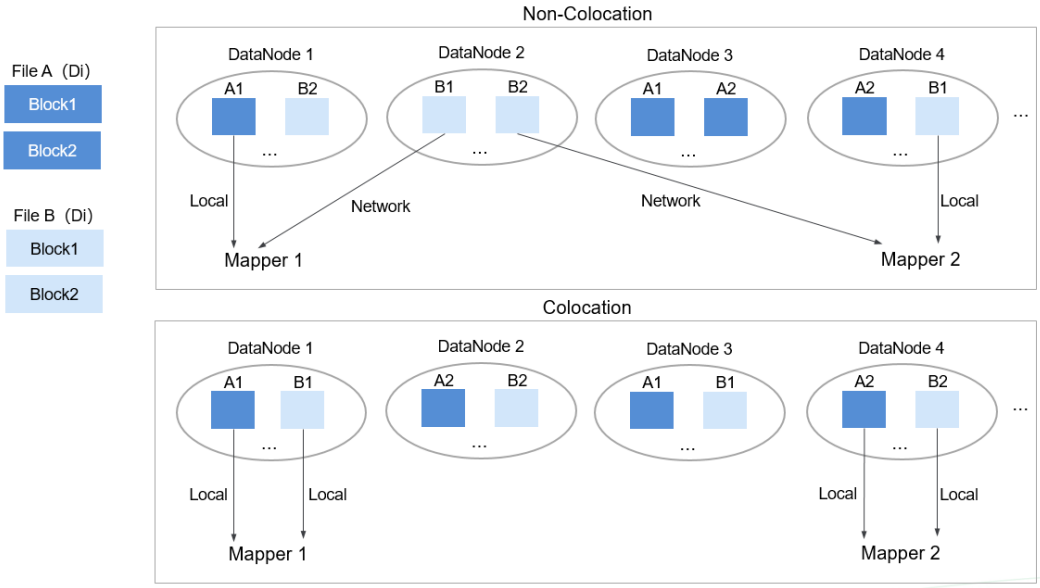


Data tables are stored in physical file system by HDFS block. Therefore, if two to-be-joined blocks are put into the same host accordingly after they are partitioned by join key, you can obtain the results directly from Map join in the local node without any data transfer in the Reduce process of the join calculation. This will greatly improve the performance.

With the identical distribution feature of HDFS data, a same distribution ID is allocated to files, FileA and FileB, on which association and summation calculations need to be performed. In this way, all the blocks are distributed together, and calculation can be performed without retrieving data across nodes, which greatly improves the MapReduce join performance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-50 Data block distribution in colocation and non-colocation scenarios

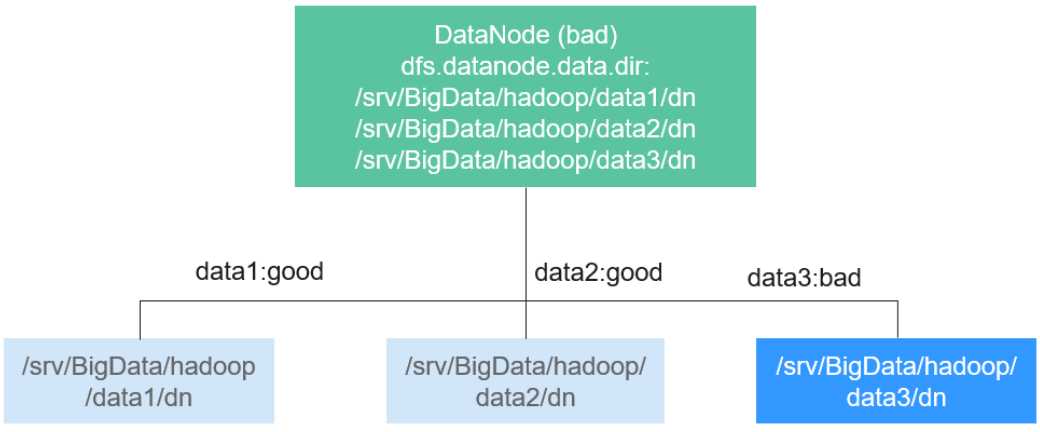


Enhanced Open Source Feature: Damaged Hard Disk Volume Configuration

In the open source version, if multiple data storage volumes are configured for a DataNode, the DataNode stops providing services by default if one of the volumes is damaged. If the configuration item **dfs.datanode.failed.volumes.tolerated** is set to specify the number of damaged volumes that are allowed, DataNode continues to provide services when the number of damaged volumes does not exceed the threshold.

The value of **dfs.datanode.failed.volumes.tolerated** ranges from -1 to the number of disk volumes configured on the DataNode. The default value is **-1**, as shown in **Figure 1-51**.

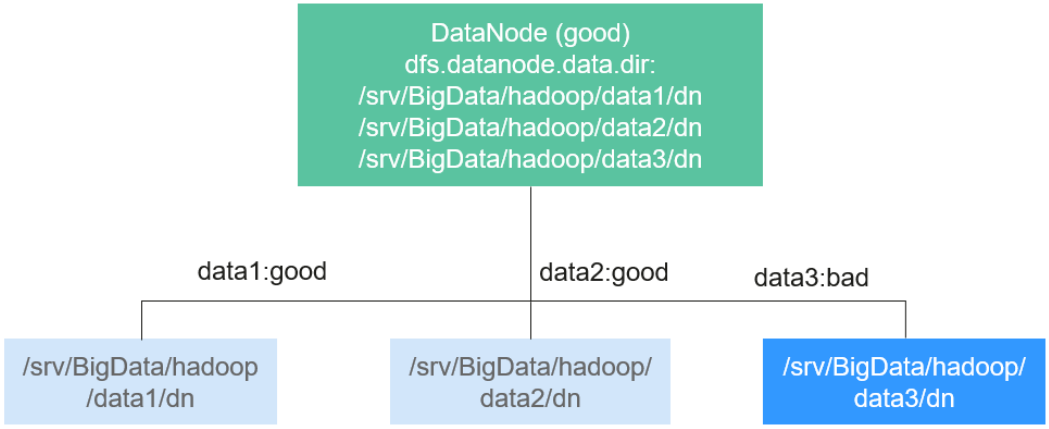
Figure 1-51 Item being set to 0



For example, three data storage volumes are mounted to a DataNode, and **dfs.datanode.failed.volumes.tolerated** is set to 1. In this case, if one data storage volume of the DataNode is unavailable, this DataNode can still provide services, as shown in **Figure 1-52**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

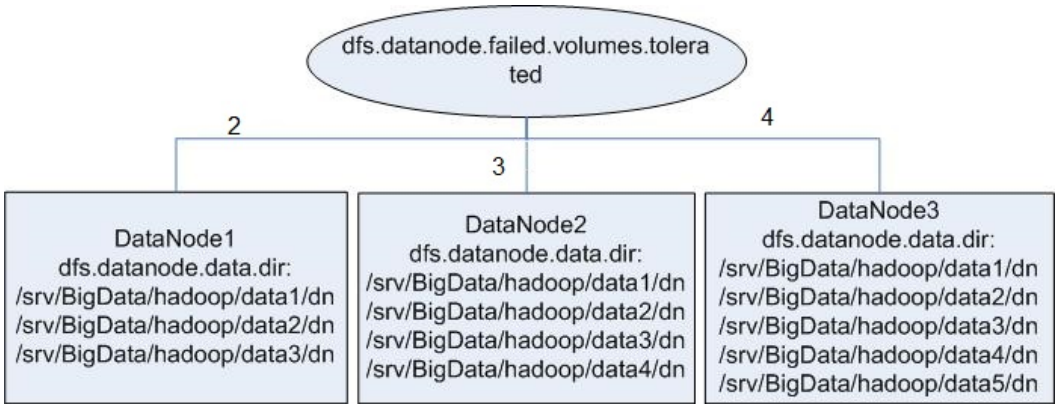
Figure 1-52 Item being set to 1



This native configuration item has some defects. When the number of data storage volumes in each DataNode is inconsistent, you need to configure each DataNode independently instead of generating the unified configuration file for all nodes.

Assume that there are three DataNodes in a cluster. The first node has three data directories, the second node has four, and the third node has five. If you want to ensure that DataNode services are available when only one data directory is available, you need to perform the configuration as shown in **Figure 1-53**.

Figure 1-53 Attribute configuration before being enhanced

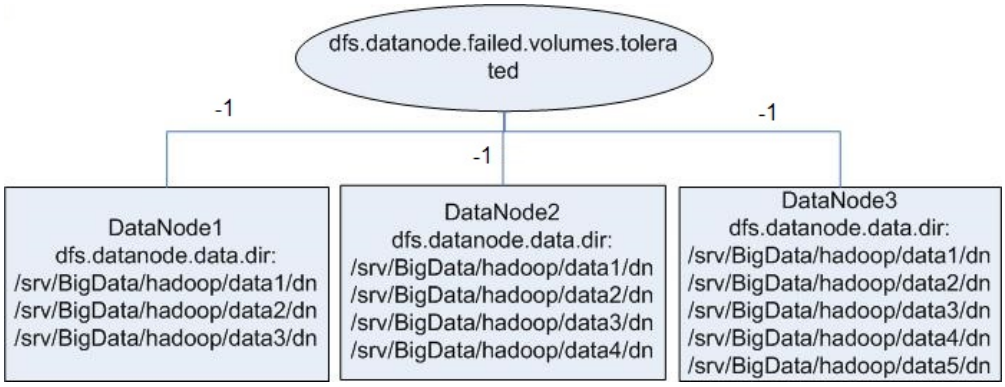


In self-developed enhanced HDFS, this configuration item is enhanced, with a value **-1** added. When this configuration item is set to **-1**, all DataNodes can provide services as long as one data storage volume in all DataNodes is available.

To resolve the problem in the preceding example, set this configuration to **-1**, as shown in **Figure 1-54**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-54 Attribute configuration after being enhanced



Enhanced Open Source Feature: HDFS Startup Acceleration

In HDFS, when NameNodes start, the metadata file FsImage needs to be loaded. Then, DataNodes will report the data block information after the DataNodes startup. When the data block information reported by DataNodes reaches the preset percentage, NameNodes exits safe mode to complete the startup process. If the number of files stored on the HDFS reaches the million or billion level, the two processes are time-consuming and will lead to a long startup time of the NameNode. Therefore, this version optimizes the process of loading metadata file FsImage.

In the open source HDFS, FsImage stores all types of metadata information. Each type of metadata information (such as file metadata information and folder metadata information) is stored in a section block, respectively. These section blocks are loaded in serial mode during startup. If a large number of files and folders are stored on the HDFS, loading of the two sections is time-consuming, prolonging the HDFS startup time. HDFS NameNode divides each type of metadata by segments and stores the data in multiple sections when generating the FsImage files. When the NameNodes start, sections are loaded in parallel mode. This accelerates the HDFS startup.

Enhanced Open Source Feature: Label-based Block Placement Policies (HDFS Nodelabel)

You need to configure the nodes for storing HDFS file data blocks based on data features. You can configure a label expression to an HDFS directory or file and assign one or more labels to a DataNode so that file data blocks can be stored on specified DataNodes. If the label-based data block placement policy is used for selecting DataNodes to store the specified files, the DataNode range is specified based on the label expression. Then proper nodes are selected from the specified range.

- You can store the replicas of data blocks to the nodes with different labels accordingly. For example, store two replicas of the data block to the node labeled with L1, and store other replicas of the data block to the nodes labeled with L2.
- You can set the policy in case of block placement failure, for example, select a node from all nodes randomly.

Figure 1-55 gives an example:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Data in **/HBase** is stored in A, B, and D.
- Data in **/Spark** is stored in A, B, D, E, and F.
- Data in **/user** is stored in C, D, and F.
- Data in **/user/shl** is stored in A, E, and F.

Figure 1-55 Example of label-based block placement policy



Enhanced Open Source Feature: HDFS Load Balance

The current read and write policies of HDFS are mainly for local optimization without considering the actual load of nodes or disks. Based on I/O loads of different nodes, the load balance of HDFS ensures that when read and write operations are performed on the HDFS client, the node with low I/O load is selected to perform such operations to balance I/O load and fully utilize the overall throughput of the cluster.

If HDFS Load Balance is enabled during file writing, the NameNode selects a DataNode (in the order of local node, local rack, and remote rack). If the I/O load of the selected node is heavy, the NameNode will choose another DataNode with lighter load.

If HDFS Load Balance is enabled during file reading, an HDFS client sends a request to the NameNode to provide the list of DataNodes that store the block to be read. The NameNode returns a list of DataNodes sorted by distance in the network topology. With the HDFS Load Balance feature, the DataNodes on the list

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

are also sorted by their I/O load. The DataNodes with heavy load are at the bottom of the list.

Enhanced Open Source Feature: HDFS Auto Data Movement

Hadoop has been used for batch processing of immense data in a long time. The existing HDFS model is used to fit the needs of batch processing applications very well because such applications focus more on throughput than delay.

However, as Hadoop is increasingly used for upper-layer applications that demand frequent random I/O access such as Hive and HBase, low latency disks such as solid state disk (SSD) are favored in delay-sensitive scenarios. To cater to the trend, HDFS supports a variety of storage types. Users can choose a storage type according to their needs.

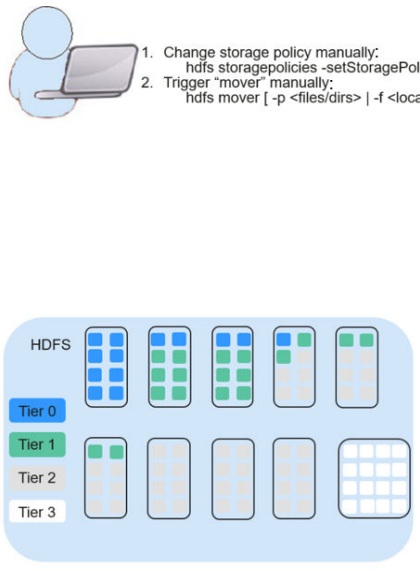
Storage policies vary depending on how frequently data is used. For example, if data that is frequently accessed in the HDFS is marked as **ALL_SSD** or **HOT**, the data that is accessed several times may be marked as **WARM**, and data that is rarely accessed (only once or twice access) can be marked as **COLD**. You can select different data storage policies based on the data access frequency.



However, low latency disks are far more expensive than spinning disks. Data typically sees heavy initial usage with decline in usage over a period of time. Therefore, it can be useful if data that is no longer used is moved out from expensive disks to cheaper ones storage media.

A typical example is storage of detail records. New detail records are imported into SSD because they are frequently queried by upper-layer applications. As access frequency to these detail records declines, they are moved to cheaper storage.

Before automatic data movement is achieved, you have to manually determine by service type whether data is frequently used, manually set a data storage policy, and manually trigger the HDFS Auto Data Movement Tool, as shown in the figure below.



Policy ID	PolicyName	Block Placement (n replicas)	Fallback storages for creation	Fallback storages for replication
15	Lazy_Persist	RAN_DISK:1 DISK:n-1	DISK	DISK
12	All_SSD	SSD:n	DISK	DISK
10	One_SSD	SSD:1,DISK:n-1	SSD,DISK	SSD,DISK
7	Hot(default)	DISK:n	<none>	ARCHIVE
5	Warm	DISK:1,ARCHIVE:n-1	ARCHIVE, DISK	ARCHIVE, DISK
2	Cold	ARCHIVE:n	<none>	<none>

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If aged data can be automatically identified and moved to cheaper storage (such as disk/archive), you will see significant cost cuts and data management efficiency improvement.

The HDFS Auto Data Movement Tool is at the core of HDFS Auto Data Movement. It automatically sets a storage policy depending on how frequently data is used. Specifically, functions of the HDFS Auto Data Movement Tool can:

- Mark a data storage policy as **All_SSD**, **One_SSD**, **Hot**, **Warm**, **Cold**, or **FROZEN** according to age, access time, and manual data movement rules.
- Define rules for distinguishing cold and hot data based on the data age, access time, and manual migration rules.
- Define the action to be taken if age-based rules are met.
  - **MARK**: the action for identifying whether data is frequently or rarely used based on the age rules and setting a data storage policy.
  - **MOVE**: the action for identifying data is frequently or rarely used based on the age rules, setting a data storage policy, triggering the HDFS Auto Data Movement tool to migrate data, and invoking the HDFS cold and hot data migration tool to migrate data across layers.
  - **SET_REPL**: the action for setting new replica quantity for a file.
  - **MOVE_TO_FOLDER**: the action for moving files to a target folder.
  - **DELETE**: the action for deleting a file or directory.
  - **SET_NODE_LABEL**: the action for setting node labels of a file.

With the HDFS Auto Data Movement feature, you only need to define age based on access time rules. HDFS Auto Data Movement Tool matches data according to age-based rules, sets storage policies, and moves data. In this way, data management efficiency and cluster resource efficiency are improved.

1.6.11 HetuEngine

1.6.11.1 HetuEngine Basic Principles

HetuEngine Description

HetuEngine is an enterprise-class high-performance interactive engine for SQL analysis and data virtualization. It seamlessly integrates with the big data ecosystem to implement interactive query of massive amounts of data within seconds, and supports cross-source and cross-domain unified data access to enable one-stop SQL convergence analysis in the data lake, between lakes, and between lakehouses.

Building on the open-source Trino SQL syntax, HetuEngine has made enterprise-level improvements in security, high availability, and ease-of-use. Its kernel capabilities continue to evolve with the open-source community.

HetuEngine Architecture

HetuEngine consists of different modules. **Figure 1-56** shows the architecture.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-56 HetuEngine architecture

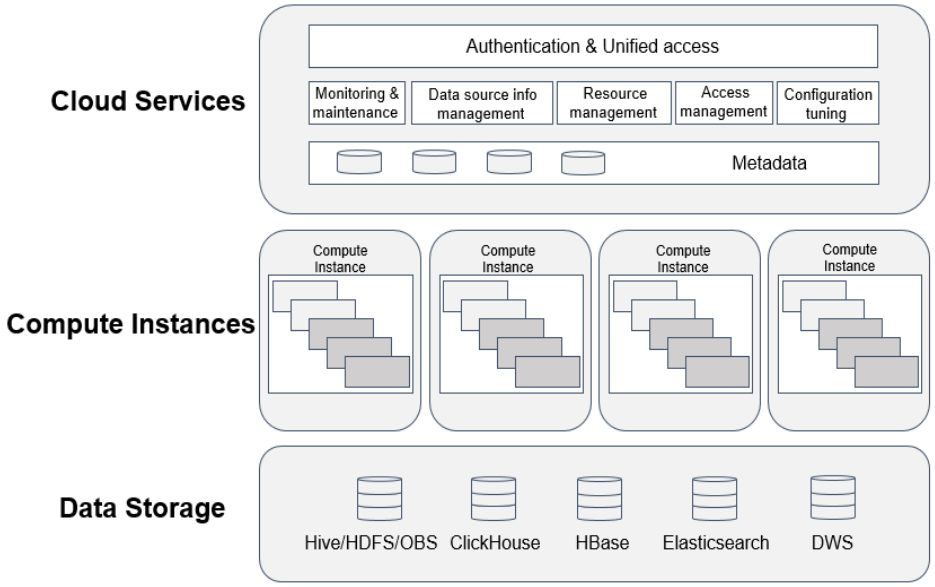


Table 1-11 Module description

Module	Concept	Description
Cloud service layer	HetuEngine CLI/ JDBC	HetuEngine client, through which the query request is submitted and the results is returned and displayed.
	HSBroker	Service management component of HetuEngine. It manages and verifies compute instances, monitors health status, and performs automatic maintenance.
	HSConsole	Provides visualized operation GUIs and RESTful APIs for data source information management, compute instance management, and automatic task query.
	HSFabric	Provides high-performance and secure data transfer across domains (data centers).
Engine layer	Coordinator	Management node of HetuEngine compute instances. It receives and parses SQL statements, generates and optimizes execution plans, assigns tasks, and schedules resources.
	Worker	Work node of HetuEngine compute instances. It provides capabilities such as parallel data pulling from data sources and distributed SQL computing.

HetuEngine Application Scenarios

HetuEngine supports cross-source (multiple data sources, such as Hive, HBase, GaussDB(DWS), and ClickHouse) and cross-domain (multiple regions or data

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

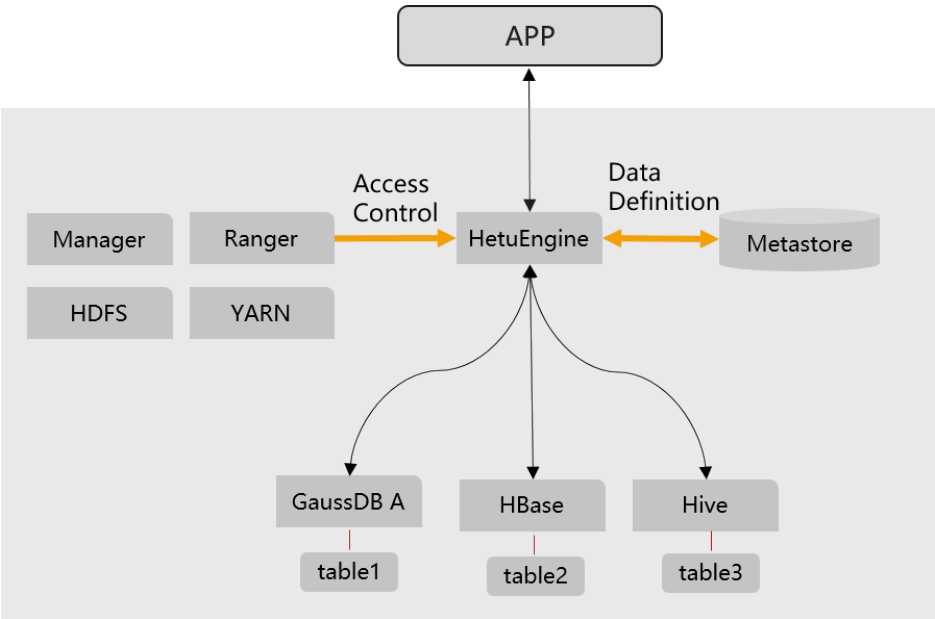
centers) quick joint query, especially for interactive quick query of Hive and Hudi data in the Hadoop cluster (MRS).

Using the HetuEngine Cross-Source Function

Enterprises usually store massive data, such as from various databases and warehouses, for management and information collection. However, diversified data sources, hybrid dataset structures, and scattered data storage rise the development cost for cross-source query and prolong the cross-source query duration.

HetuEngine provides unified standard SQL statements to implement cross-source collaborative analysis, simplifying cross-source analysis operations.

Figure 1-57 HetuEngine cross-source function

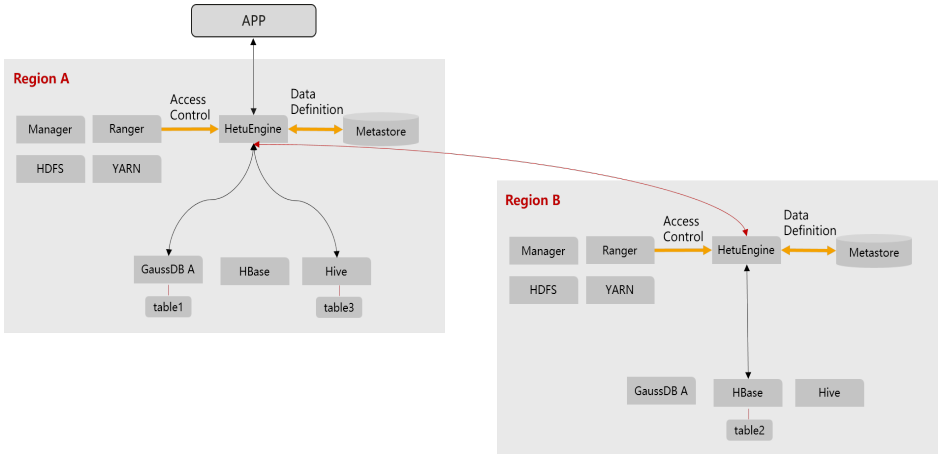


Using the HetuEngine Cross-Domain Function

HetuEngine provides unified standard SQL to implement efficient access to multiple data sources distributed in multiple regions (or data centers), shields data differences in the structure, storage, and region, and decouples data and applications.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-58 HetuEngine cross-region functions



1.6.11.2 Relationships Between HetuEngine and Other Components

The HetuEngine installation depends on the MRS cluster. **Table 1-12** lists the components on which the HetuServer installation depends.

Table 1-12 Components on which HetuEngine depends

Name	Description
HDFS	Hadoop Distributed File System, supporting high-throughput data access and suitable for applications with large-scale data sets.
Hive	Open-source data warehouse built on Hadoop. It stores structured data and implements basic data analysis using the Hive Query Language (HQL), a SQL-like language.
ZooKeeper	Enables highly reliable distributed coordination. It helps prevent single point of failures (SPOFs) and provides reliable services for applications.
KrbServer	Key management center that distributes bills.
Yarn	Resource management system, which is a general resource module that manages and schedules resources for various applications.
DBService	DBService is a high-availability relational database storage system that provides metadata backup and restoration functions.

1.6.12 Hive

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.12.1 Hive Basic Principles

Hive is a data warehouse infrastructure built on top of Hadoop. It provides a series of tools that can be used to extract, transform, and load (ETL) data. Hive is a mechanism that can store, query, and analyze mass data stored on Hadoop. Hive defines a simple SQL-like query language, which is known as Hive SQL langage (HQL). It allows a user familiar with SQL to query data. Hive data computing depends on MapReduce, Spark, and Tez.

The new execution engine Tez is used to replace the original MapReduce, significantly improving performance. Tez can convert multiple dependent jobs into one job, so only once HDFS write is required and fewer transit nodes are needed, greatly improving the performance of DAG jobs.

Hive provides the following functions:

- Analyzes massive structured data and summarizes analysis results.
- Allows complex MapReduce jobs to be compiled in SQL languages.
- Supports flexible data storage formats, including JavaScript object notation (JSON), comma separated values (CSV), TextFile, RCFile, SequenceFile, and ORC (Optimized Row Columnar).

Hive system structure:

- User interface: CLI, Client, and web UI CLI is the most frequently-used user interface. A Hive transcript is started when CLI is started. Client refers to a Hive client, and a client user connects to the Hive Server. When entering the client mode, you need to specify the node where the Hive Server resides and start the Hive Server on this node. The web UI is used to access Hive through a browser. MRS can only access Hive in client mode.
- Metadata storage: Hive stores metadata into databases, for example, MySQL and Derby. Metadata in Hive includes a table name, table columns and partitions and their properties, table properties (indicating whether a table is an external table), and the directory where table data is stored.

Hive Architecture

Hive is a single-instance service process. It provides services by translating HQL into related MapReduce jobs or HDFS operations. **Figure 1-59** shows how Hive is connected to other components.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-59 Hive architecture

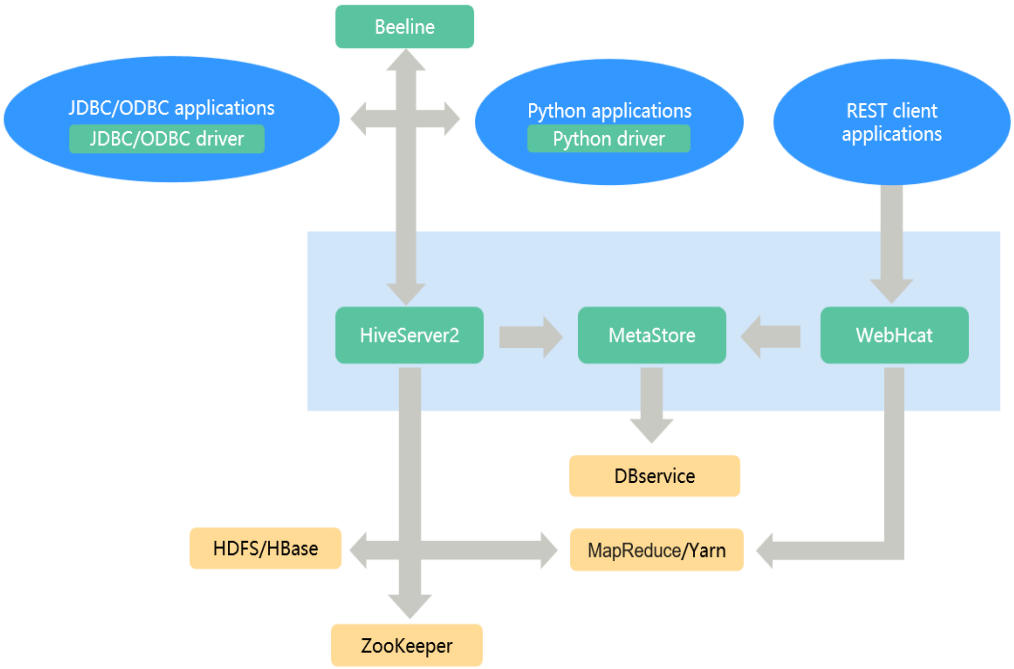


Table 1-13 Module description

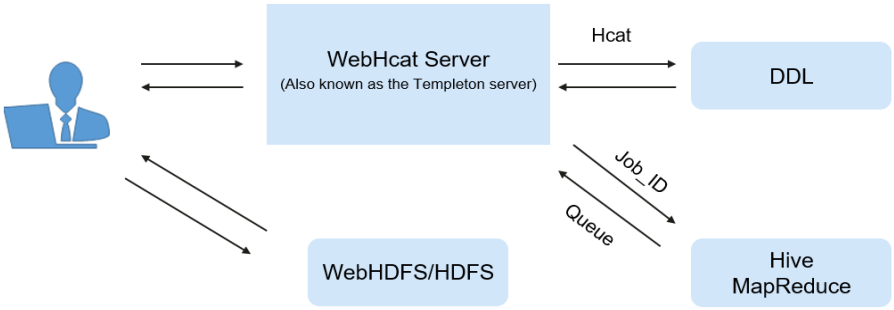
Module	Description
HiveServer	Multiple HiveServers can be deployed in a cluster to share loads. HiveServer provides Hive database services externally, translates HQL statements into related YARN tasks or HDFS operations to complete data extraction, conversion, and analysis.
MetaStore	• Multiple MetaStores can be deployed in a cluster to share loads. MetaStore provides Hive metadata services as well as reads, writes, maintains, and modifies the structure and properties of Hive tables. • MetaStore provides Thrift APIs for HiveServer, Spark, WebHCat, and other MetaStore clients to access and operate metadata.
WebHCat	Multiple WebHCats can be deployed in a cluster to share loads. WebHCat provides REST APIs and runs the Hive commands through the REST APIs to submit MapReduce jobs.
Hive client	Hive client includes the human-machine command-line interface (CLI) Beeline, JDBC drive for JDBC applications, Python driver for Python applications, and HCatalog JAR files for MapReduce.
ZooKeeper cluster	As a temporary node, ZooKeeper records the IP address list of each HiveServer instance. The client driver connects to ZooKeeper to obtain the list and selects corresponding HiveServer instances based on the routing mechanism.
HDFS/HBase cluster	The HDFS cluster stores the Hive table data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
MapReduce/ YARN cluster	It provides distributed computing services. Most Hive data operations are performed in MapReduce/YARN clusters. HiveServer can translate HQL statements into distributed computing jobs to process massive amounts of data.

HCatalog is built on Hive Metastore and incorporates the DDL capability of Hive. HCatalog is also a Hadoop-based table and storage management layer that enables convenient data read/write on tables of HDFS using different data processing tools such as MapReduce. HCatalog also provides read/write APIs for these tools and uses a Hive CLI to publish commands for defining data and querying metadata. After encapsulating these commands, WebHCat Server can provide RESTful APIs, as shown in **Figure 1-60**.

Figure 1-60 WebHCat logical architecture



Principles

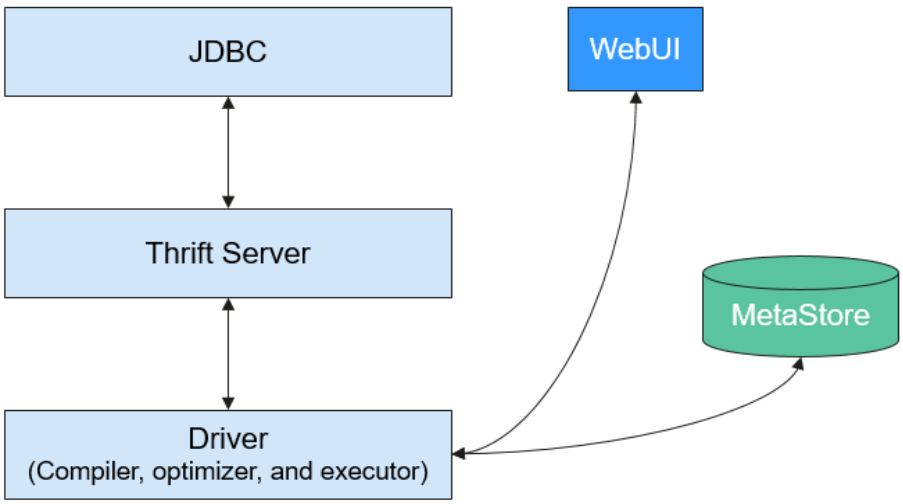
Hive functions as a data warehouse based on HDFS and MapReduce architecture and translates HQL statements into MapReduce jobs or HDFS operations.

Figure 1-61 shows the Hive structure.

- **Metastore**: reads, writes, and updates metadata such as tables, columns, and partitions. Its lower layer is relational databases.
- **Driver**: manages the lifecycle of HQL execution and participates in the entire Hive job execution.
- **Compiler**: translates HQL statements into a series of interdependent Map or Reduce jobs.
- **Optimizer**: is classified into logical optimizer and physical optimizer to optimize HQL execution plans and MapReduce jobs, respectively.
- **Executor**: runs Map or Reduce jobs based on job dependencies.
- **ThriftServer**: functions as the servers of JDBC, provides Thrift APIs, and integrates with Hive and other applications.
- **Clients**: include the web UI and JDBC APIs and provides APIs for user access.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-61 Hive framework



1.6.12.2 Hive CBO Principles

Hive CBO Principles

CBO is short for Cost-Based Optimization.

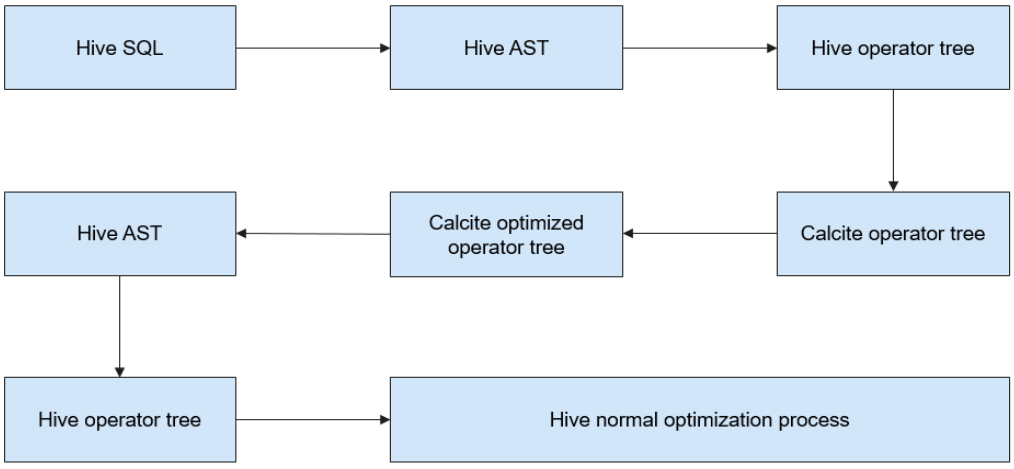
It will optimize the following:

During compilation, the CBO calculates the most efficient join sequence based on tables and query conditions involved in query statements to reduce time and resources required for query.

In Hive, the CBO is implemented as follows:

Hive uses open-source component Apache Calcite to implement the CBO. SQL statements are first converted into Hive Abstract Syntax Trees (ASTs) and then into RelNodes that can be identified by Calcite. After Calcite adjusts the join sequence in RelNodes, RelNodes are converted into ASTs by Hive to continue the logical and physical optimization. **Figure 1-62** shows the working flow.

Figure 1-62 CBO Implementation process



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Calcite adjusts the join sequence as follows:

1. A table is selected as the first table from the tables to be joined.
2. The second and third tables are selected based on the cost. In this way, multiple different execution plans are obtained.
3. A plan with the minimum costs is calculated and serves as the final sequence.

The cost calculation method is as follows:

In the current version, costs are measured based on the number of data entries after joining. Fewer data entries mean less cost. The number of joined data entries depends on the selection rate of joined tables. The number of data entries in a table is obtained based on the table-level statistics.

The number of data entries in a table after filtering is estimated based on the column-level statistics, including the maximum values (max), minimum values (min), and Number of Distinct Values (NDV).

For example, there is a table **table_a** whose total number of data records is 1,000,000 and NDV is 50. The query conditions are as follows:

```
Select * from table_a where column_a='value1';
```

The estimated number of queried data entries is: 1,000,000 x 1/50 = 20,000. The selection rate is 2%.

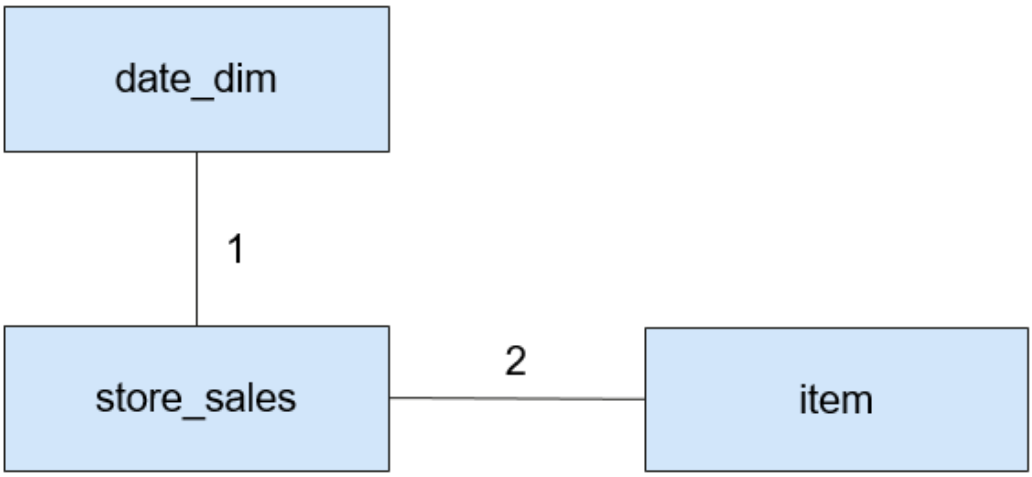
The following takes the TPC-DS Q3 as an example to describe how the CBO adjusts the join sequence:

```
select
    dt.d_year,
    item.i_brand_id brand_id,
    item.i_brand brand,
    sum(ss_ext_sales_price) sum_agg
from
    date_dim dt,
    store_sales,
    item
where
    dt.d_date_sk = store_sales.ss_sold_date_sk
    and store_sales.ss_item_sk = item.i_item_sk
    and item.i_manufact_id = 436
    and dt.d_moy = 12
group by dt.d_year , item.i_brand , item.i_brand_id
order by dt.d_year , sum_agg desc , brand_id
limit 10;
```

Statement explanation: This statement indicates that inner join is performed for three tables: table **store_sales** is a fact table with about 2,900,000,000 data entries, table **date_dim** is a dimension table with about 73,000 data entries, and table **item** is a dimension table with about 18,000 data entries. Each table has filtering conditions. **Figure 1-63** shows the join relationship.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-63 Join relationship



The CBO must first select the tables that bring better filtering effect for joining.

By analyzing min, max, NDV, and the number of data entries, the CBO estimates the selection rates of different dimension tables, as shown in **Table 1-14**.

Table 1-14 Data filtering

Table	Number of Original Data Entries	Number of Data Entries After Filtering	Selection Rate
date_dim	73,000	6,200	8.5%
item	18,000	19	0.1%

The selection rate can be estimated as follows: Selection rate = Number of data entries after filtering/Number of original data entries

As shown in the preceding table, the **item** table has a better filtering effect. Therefore, the CBO joins the **item** table first before joining the **date_dim** table.

Figure 1-64 shows the join process when the CBO is disabled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-64 Join process when the CBO is disabled

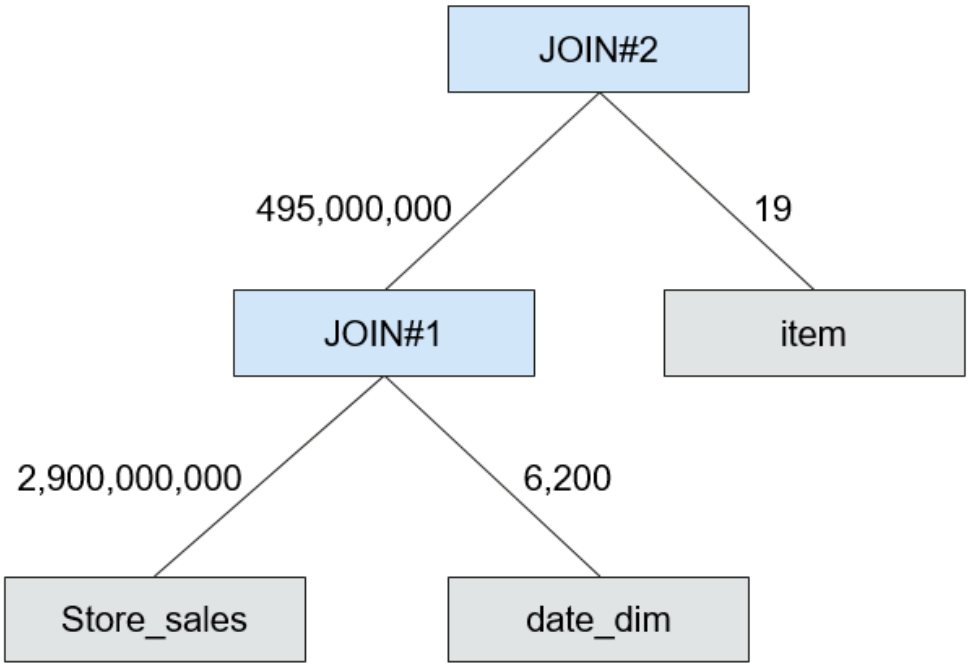
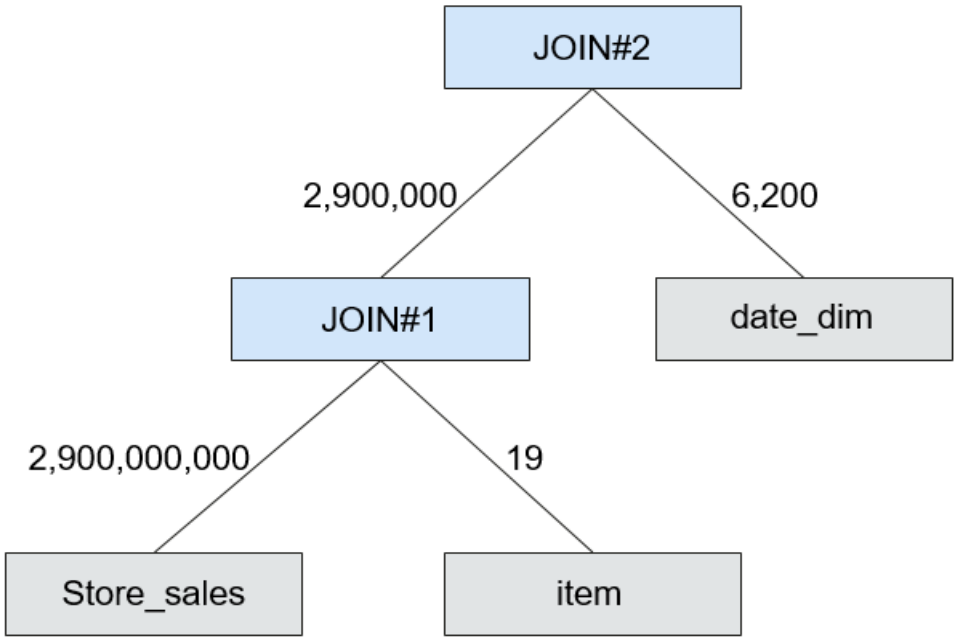


Figure 1-65 shows the join process when the CBO is enabled.

Figure 1-65 Join process when the CBO is enabled



After the CBO is enabled, the number of intermediate data entries is reduced from 495,000,000 to 2,900,000 and thus the execution time can be remarkably reduced.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.12.3 Relationship Between Hive and Other Components

HDFS

Hive is a sub-project of Apache Hadoop, which uses HDFS as the file storage system. It parses and processes structured data with highly reliable underlying storage supported by HDFS. All data files in the Hive database are stored in HDFS, and all data operations on Hive are also performed using HDFS APIs.

MapReduce

Hive data computing depends on MapReduce. MapReduce is also a sub-project of Apache Hadoop and is a parallel computing framework based on HDFS. During data analysis, Hive parses HQL statements submitted by users into MapReduce tasks and submits the tasks for MapReduce to execute.

Tez

Tez, an open-source project of Apache, is a distributed computing framework that supports directed acyclic graphs (DAGs). When Hive uses the Tez engine to analyze data, it parses HQL statements submitted by users into Tez tasks and submits the tasks to Tez for execution.

DBService

MetaStore (metadata service) of Hive processes the structure and attribute information of Hive metadata, such as Hive databases, tables, and partitions. The information needs to be stored in a relational database and is managed and processed by MetaStore. In the product, the metadata of Hive is stored and maintained by the DBService component, and the metadata service is provided by the Metadata component.

Spark

Spark can be used as the execution engine of Hive. Hive SQL statements delivered by the client are processed at the logical layer on Hive, and physical execution plans are generated and converted into a directed acyclic graph (DAG) of a resilient distributed dataset (RDD), and then submitted to a Spark cluster as a task. This way, Hive query efficiency is improved thanks to the distributed memory computing capability of Spark.

1.6.12.4 Enhanced Open Source Feature

Enhanced Open Source Feature: HDFS Colocation

HDFS Colocation is the data location control function provided by HDFS. The HDFS Colocation API stores associated data or data on which associated operations are performed on the same storage node.

Hive supports HDFS Colocation. When Hive tables are created, after the locator information is set for table files, the data files of related tables are stored on the same storage node. This ensures convenient and efficient data computing among associated tables.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Enhanced Open Source Feature: Column Encryption

Hive supports encryption of one or more columns. The columns to be encrypted and the encryption algorithm can be specified when a Hive table is created. When data is inserted into the table using the INSERT statement, the related columns are encrypted. The Hive column encryption does not support views and the Hive over HBase scenario.

The Hive column encryption mechanism supports two encryption algorithms that can be selected to meet site requirements during table creation:

- AES (the encryption class is **org.apache.hadoop.hive.serde2.AESRewriter**)
- SM4 (also known as SMS4. The encryption class is **org.apache.hadoop.hive.serde2.SMS4Rewriter**.)

Enhanced Open Source Feature: HBase Deletion

Due to the limitations of underlying storage systems, Hive does not support the ability to delete a single piece of table data. In Hive on HBase, Hive in the MRS solution supports the ability to delete a single piece of HBase table data. Using a specific syntax, Hive can delete one or more pieces of data from an HBase table.

Enhanced Open Source Feature: Row Delimiter

In most cases, a carriage return character is used as the row delimiter in Hive tables stored in text files, that is, the carriage return character is used as the terminator of a row during queries.

However, some data files are delimited by special characters, and not a carriage return character.

MRS Hive allows you to specify different characters or character combinations as row delimiters for Hive data in text files.

Enhanced Open Source Feature: HTTPS/HTTP-based REST API Switchover

WebHCat provides external REST APIs for Hive. By default, the open source community version uses the HTTP protocol.

MRS Hive supports the HTTPS protocol that is more secure, and enables switchover between the HTTP protocol and the HTTPS protocol.

Enhanced Open Source Feature: Transform Function

The Transform function is not allowed by Hive of the open source version. MRS Hive supports the configuration of the Transform function. The function is disabled by default, which is the same as that of the open source community version.

Users can modify configurations of the Transform function to enable the function. However, security risks exist when the Transform function is enabled.

Enhanced Open Source Feature: Temporary Function Creation Without ADMIN Permission

You must have **ADMIN** permission when creating temporary functions on Hive of the open source community version. MRS Hive supports the configuration of the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

function for creating temporary functions with **ADMIN** permission. The function is disabled by default, which is the same as that of the open-source community version.

You can modify configurations of this function. After the function is enabled, you can create temporary functions without **ADMIN** permission.

Enhanced Open Source Feature: Database Authorization

In the Hive open source community version, only the database owner can create tables in the database. You can be granted with the **CREATE** and **SELECT** permissions on tables by MRS Hive in a database. After you are granted with the permission to query data in the database, the system automatically associates the query permission on all tables in the database.

Enhanced Open Source Feature: Column Authorization

The Hive open source community version supports only table-level permission control. MRS Hive supports column-level permission control. You can be granted with column-level permissions, such as **SELECT**, **INSERT**, and **UPDATE**.

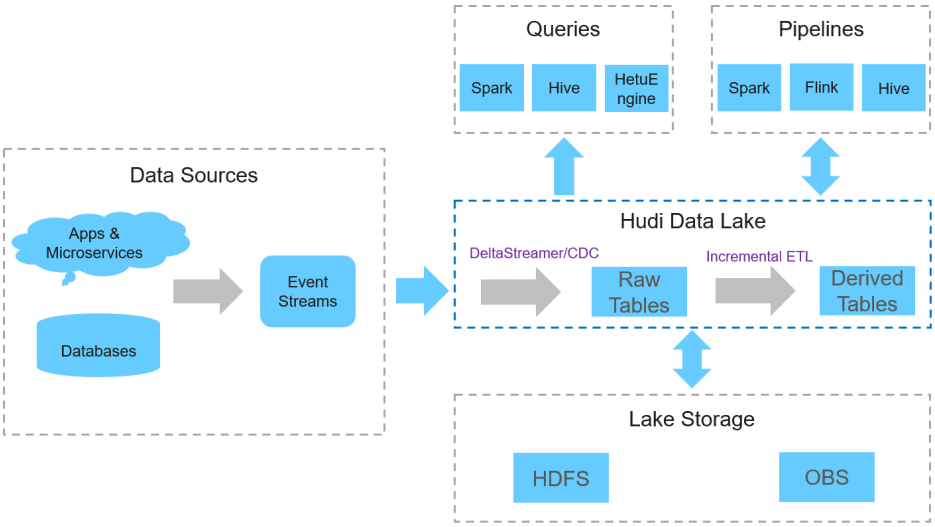
1.6.13 Hudi

Hudi is a data lake table format that provides the ability to update and delete data as well as consume new data on HDFS. It supports multiple compute engines and provides insert, update, and delete (IUD) interfaces and streaming primitives, including upsert and incremental pull, over datasets on HDFS.

NOTE

To use Hudi, ensure that the Spark/Spark2x service has been installed in the MRS cluster.

Figure 1-66 Basic architecture of Hudi



Feature

- The ACID transaction capability supports real-time data import to the lake and batch data import to the data lake.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Multiple view capabilities (read-optimized view/incremental view/real-time view) enable quick data analysis.
- Multi-version concurrency control (MVCC) design supports data version backtracking.
- Automatic management of file sizes and layouts optimizes query performance and provides quasi-real-time data for queries.
- Concurrent read and write are supported. Data can be read when being written based on snapshot isolation.
- Bootstrapping is supported to convert existing tables into Hudi datasets.

Key Technologies and Advantages

- Pluggable index mechanism: Hudi provides multiple index mechanisms to quickly update and delete massive data.
- Ecosystem support: Hudi supports multiple data engines, including Hive, Spark, HetuEngine, and Flink.

Two Types of Tables Supported by Hudi

- Copy On Write

  Copy-on-write tables are also called COW tables. Parquet files are used to store data, and internal update operations need to be performed by rewriting the original Parquet files.

  - Advantage: It is efficient because only one data file in the corresponding partition needs to be read.
  - Disadvantage: During data write, a previous copy needs to be copied and then a new data file is generated based on the previous copy. This process is time-consuming. Therefore, the data read by the read request lags behind.

- Merge On Read

  Merge-on-read tables are also called MOR tables. The combination of columnar-based Parquet and row-based format Avro is used to store data. Parquet files are used to store base data, and Avro files (also called log files) are used to store incremental data.

  - Advantage: Data is written to the delta log first, and the delta log size is small. Therefore, the write cost is low.
  - Disadvantage: Files need to be compacted periodically. Otherwise, there are a large number of fragment files. The read performance is poor because delta logs and old data files need to be merged.

Hudi Supporting Three Types Of Views for Read Capabilities in Different Scenarios

- Snapshot View

  Provides the latest snapshot data of the current Hudi table. That is, once the latest data is written to the Hudi table, the newly written data can be queried through this view.

  Both COW and MOR tables support this view capability.

- Incremental View

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Provides the incremental query capability. The incremental data after a specified commit can be queried. This view can be used to quickly pull incremental data.

COW tables support this view capability. MOR tables also support this view capability, but the incremental view capability disappears once the compact operation is performed.

- Read Optimized View

  Provides only the data stored in the latest Parquet file.

  This view is different for COW and MOR tables.

  For COW tables, the view capability is the same as the real-time view capability. (COW tables use only Parquet files to store data.)

  For MOR tables, only base files are accessed, and the data in the given file slices since the last compact operation is provided. It can be simply understood that this view provides only the data stored in Parquet files of MOR tables, and the data in log files is ignored. The data provided by this view may not be the latest. However, once the compact operation is performed on MOR tables, the incremental log data is merged into the base data. In this case, this view has the same capability as the real-time view.

1.6.14 Hue

1.6.14.1 Hue Basic Principles

Hue is a group of web applications that interact with MRS big data components. It helps you browse HDFS, perform Hive query, and start MapReduce jobs. Hue bears applications that interact with all MRS big data components.

Hue provides the file browser and query editor functions:

- File browser allows you to directly browse and operate different HDFS directories on the GUI.
- Query editor can write simple SQL statements to query data stored on Hadoop, for example, HDFS, HBase, and Hive. With the query editor, you can easily create, manage, and execute SQL statements and download the execution results as an Excel file.

On the WebUI provided by Hue, you can perform the following operations on the components:

- HDFS:
  - View, create, manage, rename, move, and delete files or directories.
  - File upload and download
  - Search for files, directories, file owners, and user groups; change the owners and permissions of the files and directories.
  - Manually configure HDFS directory storage policies and dynamic storage policies.
- Hive:
  - Edit and execute SQL/HQL statements. Save, copy, and edit the SQL/HQL template. Explain SQL/HQL statements. Save the SQL/HQL statement and query it.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Database presentation and data table presentation
  - Supporting different types of Hadoop storage
  - Use MetaStore to add, delete, modify, and query databases, tables, and views.

    NOTE

    If Internet Explorer is used to access the Hue page to execute HSQL statements, the execution fails, because the browser has functional problems. You are advised to use a compatible browser, for example, Google Chrome.

- Impala:
  - Edit and execute SQL/HQL statements. Save, copy, and edit the SQL/HQL template. Explain SQL/HQL statements. Save the SQL/HQL statement and query it.
  - Database presentation and data table presentation
  - Supporting different types of Hadoop storage
  - Use MetaStore to add, delete, modify, and query databases, tables, and views.

    NOTE

    If Internet Explorer is used to access the Hue page to execute HiveSQL statements, the execution fails, because the browser has functional problems. You are advised to use a compatible browser, for example, Google Chrome.
- MapReduce: Check MapReduce tasks that are being executed or have been finished in the clusters, including their status, start and end time, and run logs.
- Oozie: Hue provides the Oozie job manager function, in this case, you can use Oozie in GUI mode.
- ZooKeeper: Hue provides the ZooKeeper browser function for you to use ZooKeeper in GUI mode.

Architecture

Hue, adopting the MTV (Model-Template-View) design, is a web application program running on Django Python. (Django Python is a web application framework that uses open source codes.)

Hue consists of Supervisor Process and WebServer. Supervisor Process is the core Hue process that manages application processes. Supervisor Process and WebServer interact with applications on WebServer through Thrift/REST APIs, as shown in **Figure 1-67**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-67 Hue architecture

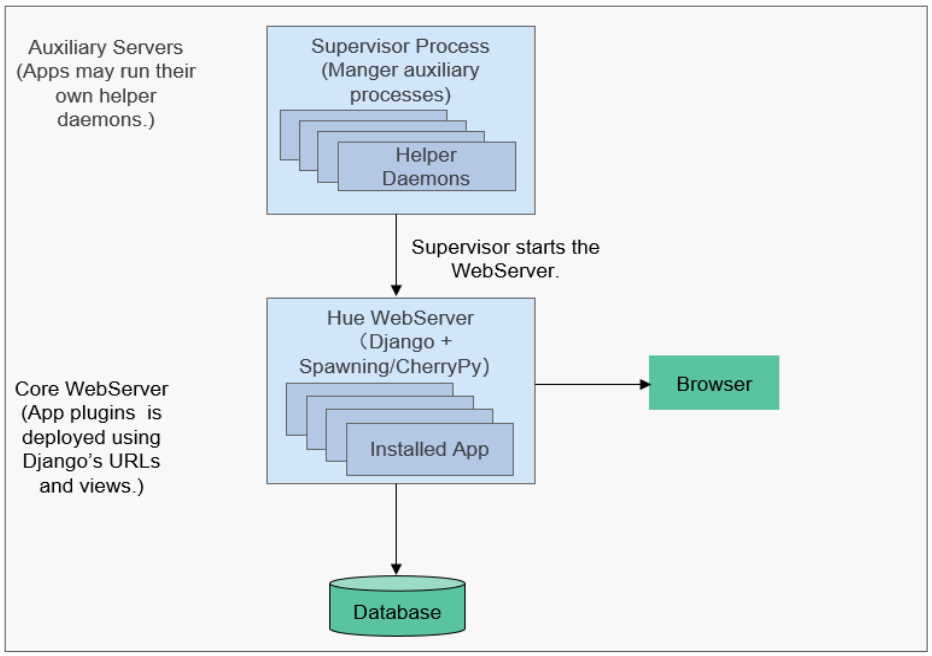


Table 1-15 describes the components shown in **Figure 1-67**.

Table 1-15 Architecture description

Connection Name	Description
Supervisor Process	Manages processes of WebServer applications, such as starting, stopping, and monitoring the processes.
Hue WebServer	Provides the following functions through the Django Python web framework: • Deploys applications. • Provides the GUI. • Connects to databases to store persistent data of applications.

1.6.14.2 Relationship Between Hue and Other Components

Relationship Between Hue and Hadoop Clusters

Figure 1-68 shows how Hue interacts with Hadoop clusters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-68 Hue and Hadoop clusters

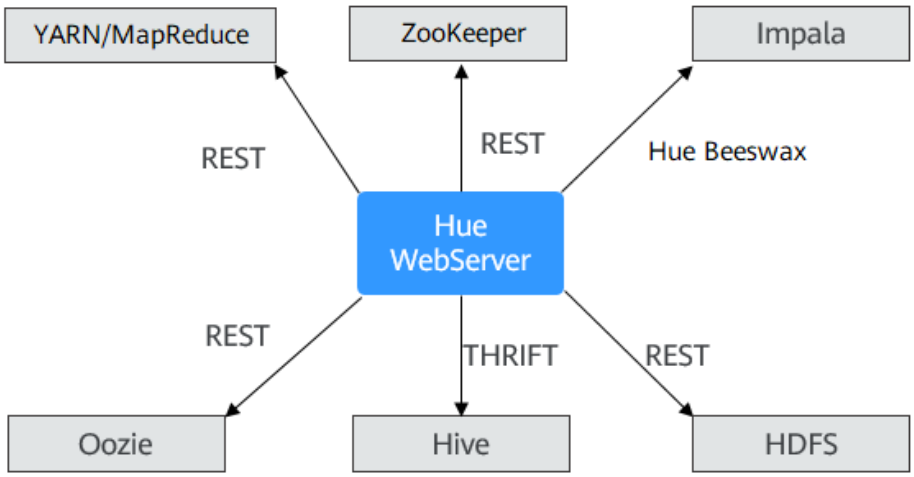


Table 1-16 Relationships between Hue and other components

Connection Name	Description
HDFS	HDFS provides REST APIs to interact with Hue to query and operate HDFS files. Hue packages a user request into interface data, sends the request to HDFS through REST APIs, and displays execution results on the web UI.
Hive	Hive provides Thrift interfaces to interact with Hue, execute Hive SQL statements, and query table metadata. If you edit HQL statements on the Hue web UI, then, Hue submits the HQL statements to the Hive server through the Thrift APIs and displays execution results on the web UI.
YARN/ MapReduce	MapReduce provides REST APIs to interact with Hue and query YARN job information. If you go to the Hue web UI, enter the filter parameters, the UI sends the parameters to the background, and Hue invokes the REST APIs provided by MapReduce (MR1/MR2-YARN) to obtain information such as the status of the task running, the start/end time, the run log, and more.
Oozie	Oozie provides REST APIs to interact with Hue, create workflows, coordinators, and bundles, and manage and monitor tasks. A graphical workflow, coordinator, and bundle editor are provided on the Hue web UI. Hue invokes the REST APIs of Oozie to create, modify, delete, submit, and monitor workflows, coordinators, and bundles.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Connection Name	Description
ZooKeeper	ZooKeeper provides REST APIs to interact with Hue and query ZooKeeper node information. ZooKeeper node information is displayed in the Hue web UI. Hue invokes the REST APIs of ZooKeeper to obtain the node information.
Impala	Impala provides Hue Beeswax APIs to interact with Hue, execute Hive SQL statements, and query table metadata. If you edit HQL statements on the Hue web UI, then, Hue submits the HQL statements to the Hive server through the Hue Beeswax APIs and displays execution results on the web UI.

1.6.14.3 Hue Enhanced Open Source Features

Hue Enhanced Open Source Features

- Storage policy: The number of HDFS file copies varies depending on the storage media. This feature allows you to manually set an HDFS directory storage policy or can automatically adjust the file storage policy, modify the number of file copies, move the file directory, and delete files based on the latest access time and modification time of HDFS files to fully utilize storage capacity and improve storage performance.
- MR engine: You can use the MapReduce engine to execute Hive SQL statements.
- Reliability enhancement: Hue is deployed in active/standby mode. When interconnecting with HDFS, Oozie, Hive, and YARN, Hue can work in failover or load balancing mode.

1.6.15 Impala

Impala

Impala provides fast, interactive SQL queries directly on your Apache Hadoop data stored in HDFS, HBase, or the Object Storage Service (OBS). Impala uses the same unified storage platform, metadata, SQL syntax (Hive SQL), ODBC driver, and user interface (Impala query UI in Hue) as Apache Hive. This provides a familiar and unified platform for real-time or batch-oriented queries. Impala is an addition to tools available for querying big data. Impala does not replace the batch processing frameworks built on MapReduce such as Hive. Hive and other frameworks built on MapReduce are best suited for long running batch jobs.

Impala provides the following features:

- Most common SQL-92 features of Hive Query Language (HQL) including SELECT, JOIN, and aggregate functions
- HDFS, HBase, and OBS storage, including:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - HDFS file formats: delimited text files, Parquet, Avro, SequenceFile, and RCFile
  - Compression codecs: Snappy, GZIP, Deflate, BZIP
- Common data access interfaces including:
  - JDBC driver
  - ODBC driver
  - Hue Beeswax and the Impala query UI
- **Impala-shell** command line interface
- Kerberos authentication

Impala applies to offline analysis (such as log and cluster status analysis) of real-time data queries, large-scale data mining (such as user behavior analysis, interest region analysis, and region display), and other scenarios.

Impala consists of three roles: Impala Daemon (Impalad), Impala StateStore, and Impala Catalog Service.

Impala Daemon

The core Impala component is the Impala daemon, physically represented by the **impalad** process.

A few of the key functions that an Impala daemon performs are:

- Runs on all data nodes.
- Reads and writes to data files.
- Accepts queries transmitted from the **Impala-shell** command, Hue, JDBC, or ODBC.
- Parallelizes the queries and transmits intermediate query results back to the central coordinator.
- Invokes a node to return the query results to the client.

The Impala daemons are in constant communication with StateStore, to confirm which daemons are healthy and can accept new work.

Impala StateStore

The Impala component known as the StateStore checks on the health of all Impala daemons in a cluster, and continuously relays its findings to each of those daemons. It is physically represented by a daemon process named **statestored**. You only need such a process on one host in a cluster. If an Impala daemon goes offline due to hardware failure, network error, software issue, or other reason, the StateStore informs all the other Impala daemons so that future queries can avoid making requests to the unreachable Impala daemon.

Impala Catalog Service

The Impala component known as the Catalog Service relays the metadata changes from Impala SQL statements to all the Impala daemons in a cluster. It is physically represented by a daemon process named **catalogd**. When you create a table, load data, and so on through Hive, you do need to issue REFRESH or INVALIDATE

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

METADATA on an Impala daemon before executing a query there. The catalog service avoids the need to issue REFRESH and INVALIDATE METADATA statements when the metadata changes are performed by statements issued through Impala.

Relationship with Other Components

- Hadoop Distributed File System (HDFS)

  Impala uses the HDFS to store files. It parses and processes structured data with highly reliable underlying storage supported by HDFS. Impala does not move data in HDFS and provides faster access.

- Hive

  Impala uses Hive metadata, Open Database Connectivity (ODBC) driver, and SQL syntax. Unlike Hive, which is over MapReduce, Impala implements a distributed architecture based on daemon and handles all query executions on the same node. Therefore, Impala is faster than Hive by reducing the latency caused by MapReduce.

- Kudu

  Kudu is closely integrated with Impala to replace the combination of Impala, HDFS, and Parquet, allowing you to insert, query, update, and delete data from Kudu tablets using Impala's SQL syntax. In addition, you can use JDBC or ODBC. Impala functions as a proxy to connect to Kudu for data operations.

- HBase

  By default, Impala tables use data files stored in HDFS to facilitate batch loading and query in full table scans. However, HBase can provide convenient and efficient query of OLTP-style organization data.

1.6.16 IoTDB

1.6.16.1 IoTDB Basic Principles

Database for Internet of Things (IoTDB) is a software system that collects, stores, manages, and analyzes IoT time series data. Apache IoTDB uses a lightweight architecture and features high performance and rich functions.

IoTDB sorts time series and stores indexes and chunks, greatly improving the query performance of time series data. IoTDB uses the Raft protocol to ensure data consistency. In time series scenarios, IoTDB pre-computes and stores data to improve analysis performance. Based on the characteristics of time series data, IoTDB provides powerful data encoding and compression capabilities. In addition, its replica mechanism ensures data security. IoTDB is deeply integrated with Apache Hadoop and Flink to meet the requirements of massive data storage, high-speed data reading, and complex data analysis in the industrial IoT field.

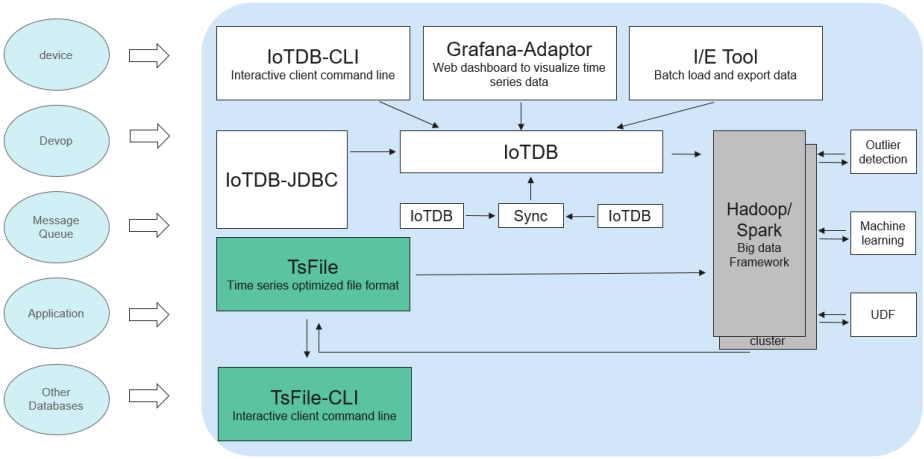
IoTDB Architecture

The IoTDB suite consists of multiple components to provide a series of functions such as data collection, data writing, data storage, data query, data visualization, and data analysis.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-69 shows the overall application architecture after all components of the IoTDB suite are used. IoTDB refers to the time series database component in the suite.

Figure 1-69 IoTDB architecture



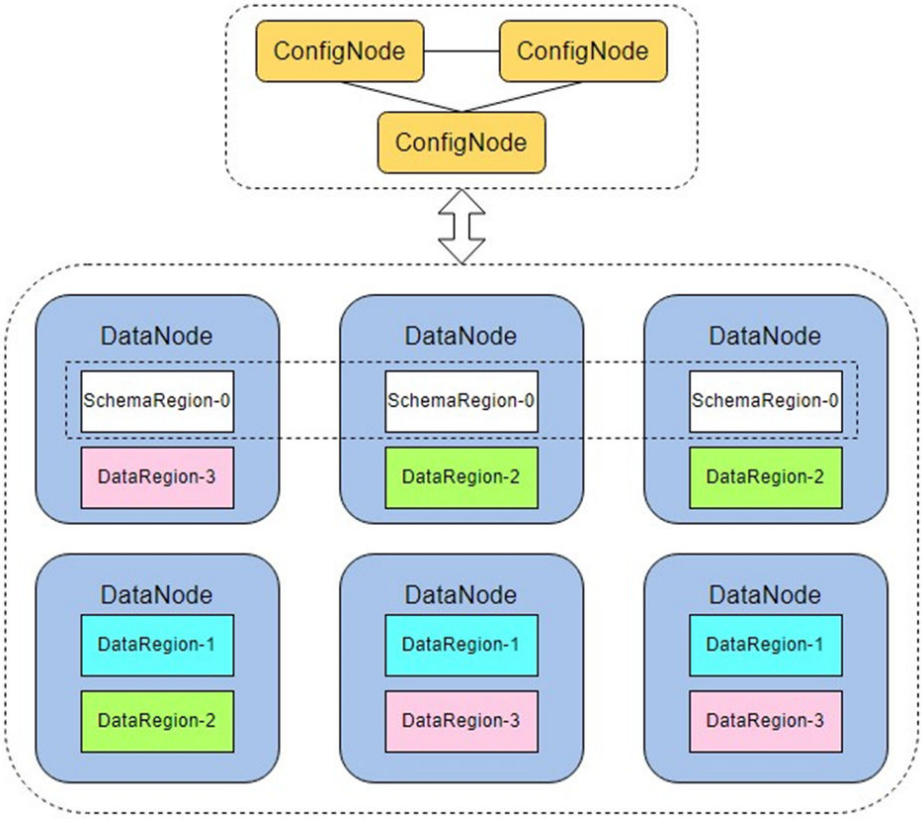
- Users can use Java Database Connectivity (JDBC) or Session to import the time series data and system status data (such as server load, CPU usage and memory usage) collected from device sensors, as well as time series data in message queues, applications, or other databases, to the local or remote IoTDB. Users can also directly write the preceding data into a local TsFile file or a TsFile file in the HDFS.
- You can write TsFile files to HDFS to meet the access requirements of data processing tasks such as Hadoop and Flink.
- The TsFile-Hadoop or TsFile-Flink connector can be used to allow Hadoop or Flink to process the TsFile files written to the HDFS or local host.
- The analysis result can be written back to a TsFile in the same way.
- IoTDB and TsFile also provide client tools to meet users' requirements for writing and viewing data in SQL, script, and graphical forms.

The IoTDB service includes two roles: IoTDBServer (DataNode) and ConfigNode. The role name DataNode of the community edition has the same name as the HDFS role. DataNode is renamed IoTDBServer.

- ConfigNode: management role, which is responsible for DataNode data sharding and load balancing.
- IoTDBServer (DataNode): storage role, which is responsible for storing, querying, and writing data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-70 IoTDB distributed architecture

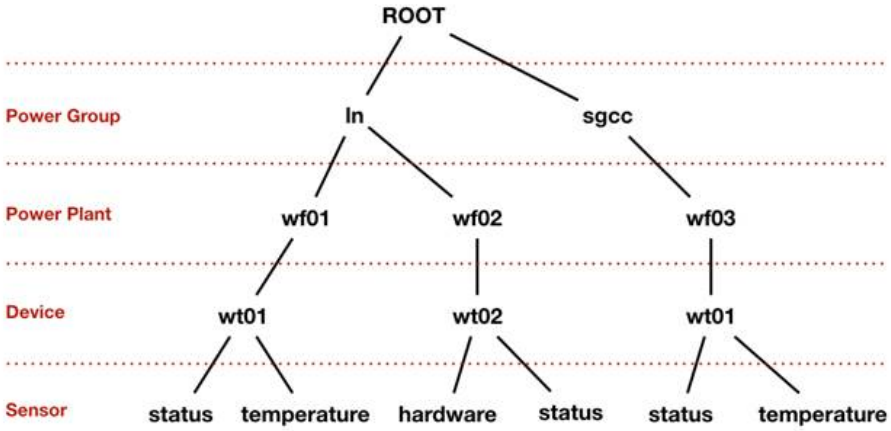


IoTDB Principles

Based on the attribute hierarchy, attribute coverage, and subordinate relationships between data, the IoTDB data model can be represented as the attribute hierarchy, as shown in **Figure 1-71**. The hierarchy is as follows: power group layer - power plant layer - device layer - sensor layer. **ROOT** is a root node, and each node at the sensor layer is a leaf node. According to the IoTDB syntax, the path from **ROOT** to a leaf node is separated by a dot (.). The complete path is used to name a time series in the IoTDB. For example, the time series name corresponding to the path on the left in the following figure is **ROOT.ln.wf01.wt01.status**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-71 IoTDB data model



Relationships Between IoTDB and Other Components

The IoTDB stores data locally, so it does not depend on any other component for storage. However, in a security cluster environment, IoTDB depends on the KrbServer component for Kerberos authentication.

1.6.16.2 IoTDB Enhanced Open Source Features

Visualization

- Visualized O&M covers installation, uninstallation, one-click start and stop, configurations, clients, monitoring, alarms, health checks, and logs.
- Visualized permission management does not require background command line operations and supports read and write permission control at the database and table levels.
- Visualized log level configuration dynamically takes effect, supports visualized download and retrieval, and supports log audit.

Security Hardening

User authentication supports Kerberos authentication and SSL encryption, which are compatible with the community authentication mode.

Ecosystem Interconnection

On the basis of native capabilities, the cluster interconnection with MQTT is enhanced.

1.6.17 JobGateway

JobGateway allows you to submit jobs through REST APIs.

As a gateway component for submitting big data jobs, JobGateway provides fully controllable enterprise-level big data job submission services that can easily submit Spark, Flink, Hive, and other big data jobs.

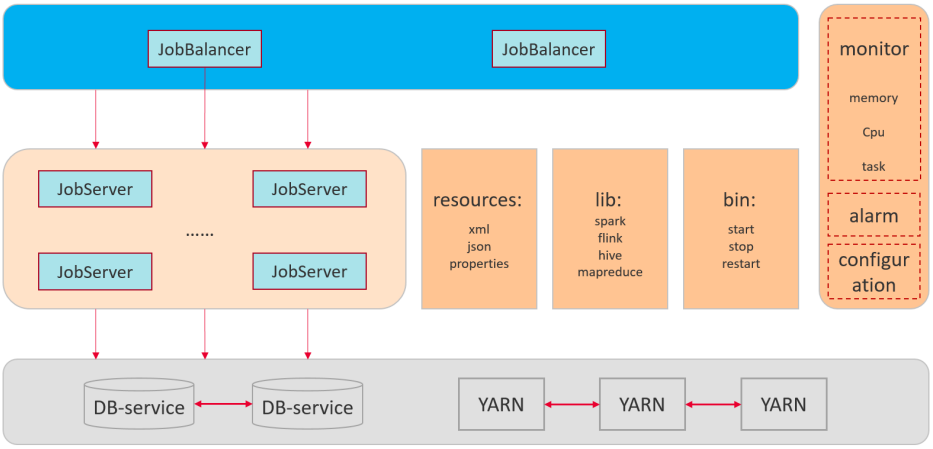
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

JobGateway Architecture

JobGateway consists of JobServer and JobBalancer instances.

- JobBalancer provides load balancing.
- JobServer provides REST APIs for submitting jobs.

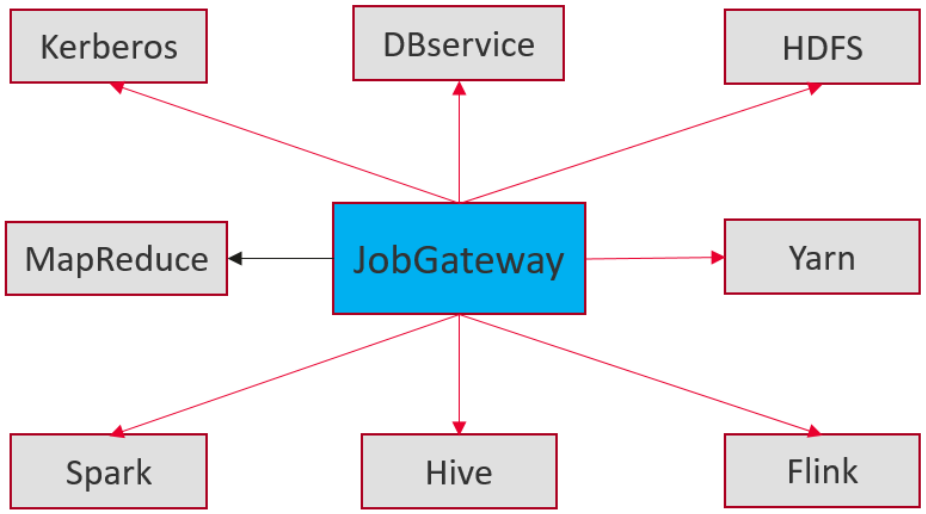
Figure 1-72 JobGateway Architecture



Relationship Between JobGateway and Other Components

JobGateway is a service that allows you to submit Spark, Hive, MapReduce, and Flink jobs through REST APIs.

Figure 1-73 Relationship between JobGateway and other components



1.6.18 Kafka

1.6.18.1 Kafka Basic Principles

Kafka is an open source, distributed, partitioned, and replicated commit log service. Kafka is publish-subscribe messaging, rethought as a distributed commit

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

log. It provides features similar to Java Message Service (JMS) but another design. It features message endurance, high throughput, distributed methods, multi-client support, and real time. It applies to both online and offline message consumption, such as regular message collection, website activeness tracking, aggregation of statistical system operation data (monitoring data), and log collection. These scenarios engage large amounts of data collection for Internet services.

Kafka Structure

Producers publish data to topics, and consumers subscribe to the topics and consume messages. A broker is a server in a Kafka cluster. For each topic, the Kafka cluster maintains partitions for scalability, parallelism, and fault tolerance. Each partition is an ordered, immutable sequence of messages that is continually appended to - a commit log. Each message in a partition is assigned a sequential ID, which is called offset.

Figure 1-74 Kafka architecture

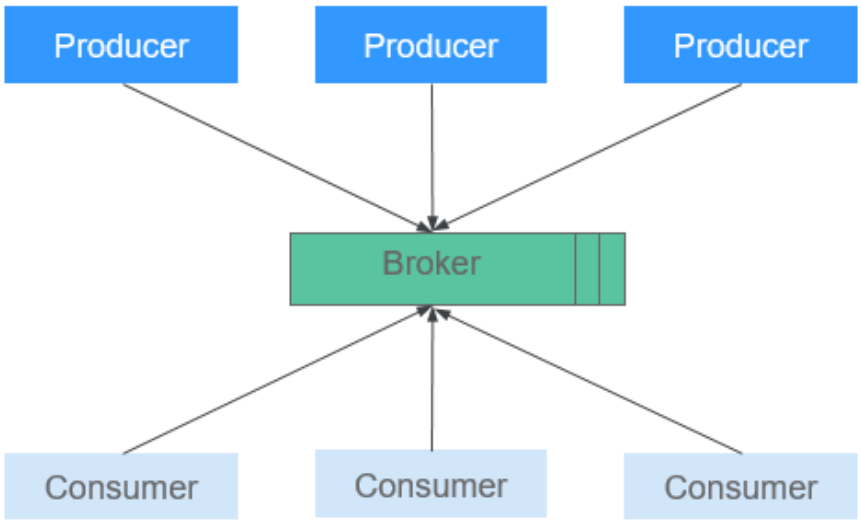


Table 1-17 Kafka architecture description

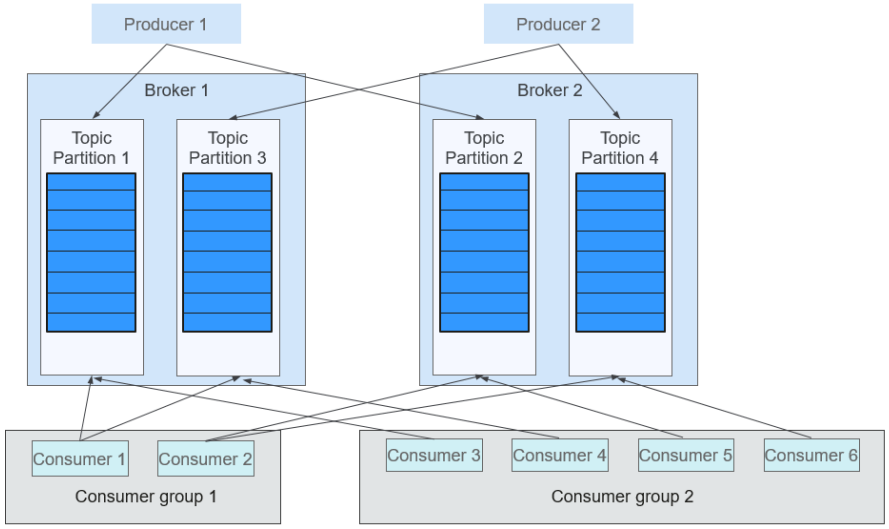
Name	Description
Broker	A broker is a server in a Kafka cluster.
Topic	A topic is a category or feed name to which messages are published. A topic can be divided into multiple partitions, which can act as a parallel unit.
Partition	A partition is an ordered, immutable sequence of messages that is continually appended to - a commit log. The messages in the partitions are each assigned a sequential ID number called the offset that uniquely identifies each message within the partition.
Producer	Producers publish messages to a Kafka topic.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Description
Consumer	Consumers subscribe to topics and process the feed of published messages.

Figure 1-75 shows the relationships between modules.

Figure 1-75 Relationships between Kafka modules



Consumers label themselves with a consumer group name, and each message published to a topic is delivered to one consumer instance within each subscribing consumer group. If all the consumer instances belong to the same consumer group, loads are evenly distributed among the consumers. As shown in the preceding figure, Consumer1 and Consumer2 work in load-sharing mode; Consumer3, Consumer4, Consumer5, and Consumer6 work in load-sharing mode. If all the consumer instances belong to different consumer groups, messages are broadcast to all consumers. As shown in the preceding figure, the messages in Topic 1 are broadcast to all consumers in Consumer Group1 and Consumer Group2.

Principle

- **Message Reliability**

  When a Kafka broker receives a message, it stores the message on a disk persistently. Each partition of a topic has multiple replicas stored on different broker nodes. If one node is faulty, the replicas on other nodes can be used.

- **High Throughput**

  Kafka provides high throughput in the following ways:

  – Messages are written into disks instead of being cached in the memory. This enables full utilization of the sequential read and write performance of disks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - The use of zero-copy eliminates I/O operations.
  - Data is sent in batches, improving network utilization.
  - Each topic is divided in to multiple partitions, which increases concurrent processing. Concurrent read and write operations can be performed between multiple producers and consumers. Producers send messages to specified partitions based on the algorithm used.
- **Message Subscribe-Notify Mechanism**

  Consumers subscribe to interested topics and consume data in pull mode. Consumers can choose the consumption mode, such as batch consumption, repeated consumption, and consumption from the end, and control the message pulling speed based on actual situation. Consumers need to maintain the consumption records by themselves.
- **Scalability**

  When broker nodes are added to a Kafka cluster, they are registered with ZooKeeper. After the registration is successful, procedures and consumers can sense the change in a timely manner and make related adjustment.

Open Source Features

- Reliability

  Message processing methods such as **At-Least Once**, **At-Most Once**, and **Exactly Once** are provided. The message processing status is maintained by consumers. Kafka needs to work with the application layer to implement **Exactly Once**.
- High throughput

  High throughput is provided for message publishing and subscription.
- Persistence

  Messages are stored on disks and can be used for batch consumption and real-time application programs. Data persistence and replication prevent data loss.
- Distribution

  A distributed system is easy to be expanded externally. Multiple Producers, Brokers, and Consumers can be deployed in each cluster to form a distributed cluster. The system can be expanded without shutting down the system.

1.6.18.2 Relationship Between Kafka and Other Components

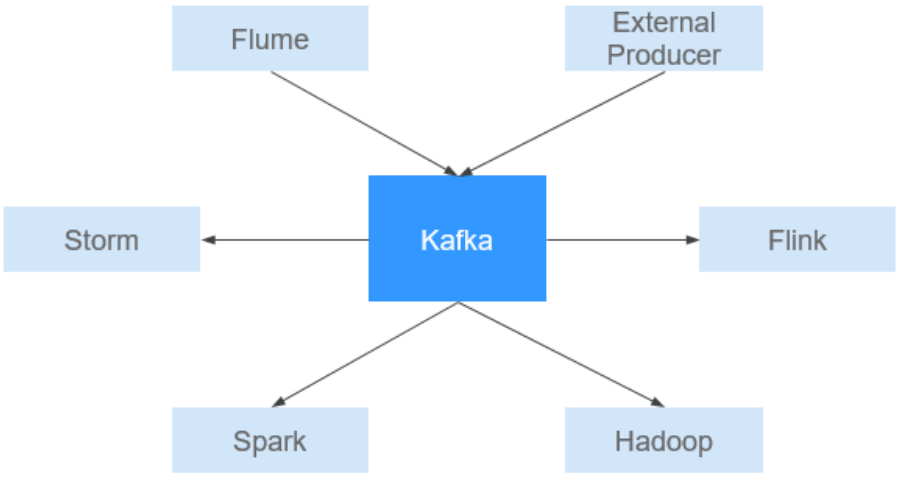
As a message publishing and subscription system, Kafka allows high-speed data transmission between different subsystems of the FusionInsight platform.

It can receive external messages in real time and send the messages to online and offline services for processing.

The following figure shows the relationship between Kafka and other components.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-76 Relationship with other components



1.6.18.3 Kafka Enhanced Open Source Features

Kafka Enhanced Open Source Features

- Monitors the following topic-level metrics:
  - Topic Input Traffic
  - Topic Output Traffic
  - Topic Rejected Traffic
  - Number of Failed Fetch Requests Per Second
  - Number of Failed Produce Requests Per Second
  - Number of Topic Input Messages Per Second
  - Number of Fetch Requests Per Second
  - Number of Produce Requests Per Second
- Queries the mapping between broker IDs and node IP addresses. On Linux clients, **kafka-broker-info.sh** can be used to query the mapping between broker IDs and node IP addresses.

1.6.19 KrbServer and LdapServer

1.6.19.1 KrbServer and LdapServer Principles

Overview

To manage the access control permissions on data and resources in a cluster, it is recommended that the cluster in security mode be installed. In security mode, a client application must be authenticated and a secure session must be established before the application accesses any resource in the cluster. MRS uses KrbServer to provide Kerberos authentication for all components, implementing a reliable authentication mechanism.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

LdapServer supports Lightweight Directory Access Protocol (LDAP) and provides the capability of storing user and user group data for Kerberos authentication.

Architecture

The security authentication function for user login depends on Kerberos and LDAP.

Figure 1-77 Security authentication architecture

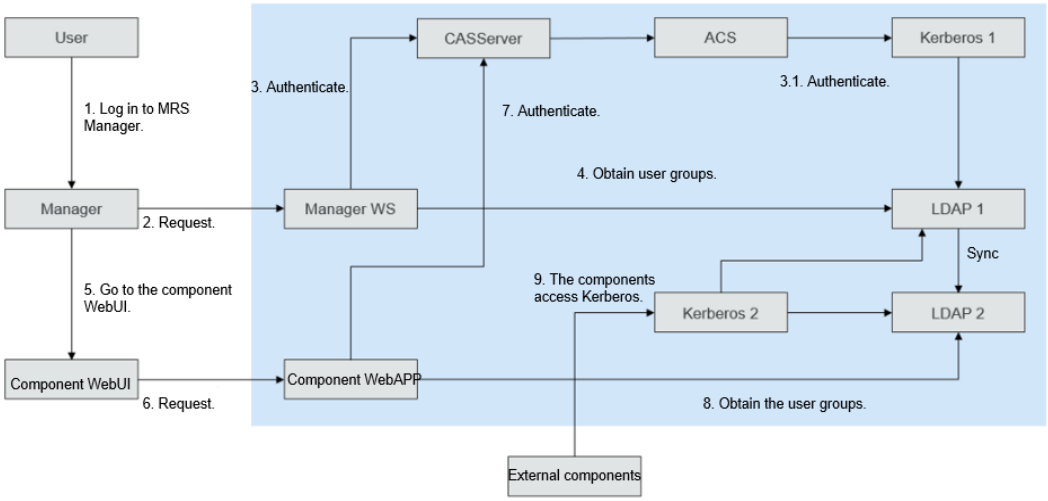


Figure 1-77 includes three scenarios:

- Logging in to the MRS Manager Web UI
  The authentication architecture includes steps 1, 2, 3, and 4.
- Logging in to a component web UI
  The authentication architecture includes steps 5, 6, 7, and 8.
- Accessing between components
  The authentication architecture includes step 9.

Table 1-18 Key modules

Connection Name	Description
Manager	Cluster Manager
Manager WS	WebBrowser
Kerberos1	KrbServer (management plane) service deployed in MRS Manager, that is, OMS Kerberos
Kerberos2	KrbServer (service plane) service deployed in the cluster
LDAP1	LdapServer (management plane) service deployed in MRS Manager, that is, OMS LDAP
LDAP2	LdapServer (service plane) service deployed in the cluster

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Data operation mode of Kerberos1 in LDAP: The active and standby instances of LDAP1 and the two standby instances of LDAP2 can be accessed in load balancing mode. Data write operations can be performed only in the active LDAP1 instance. Data read operations can be performed in LDAP1 or LDAP2.

Data operation mode of Kerberos2 in LDAP: Data read operations can be performed in LDAP1 and LDAP2. Data write operations can be performed only in the active LDAP1 instance.

Principle

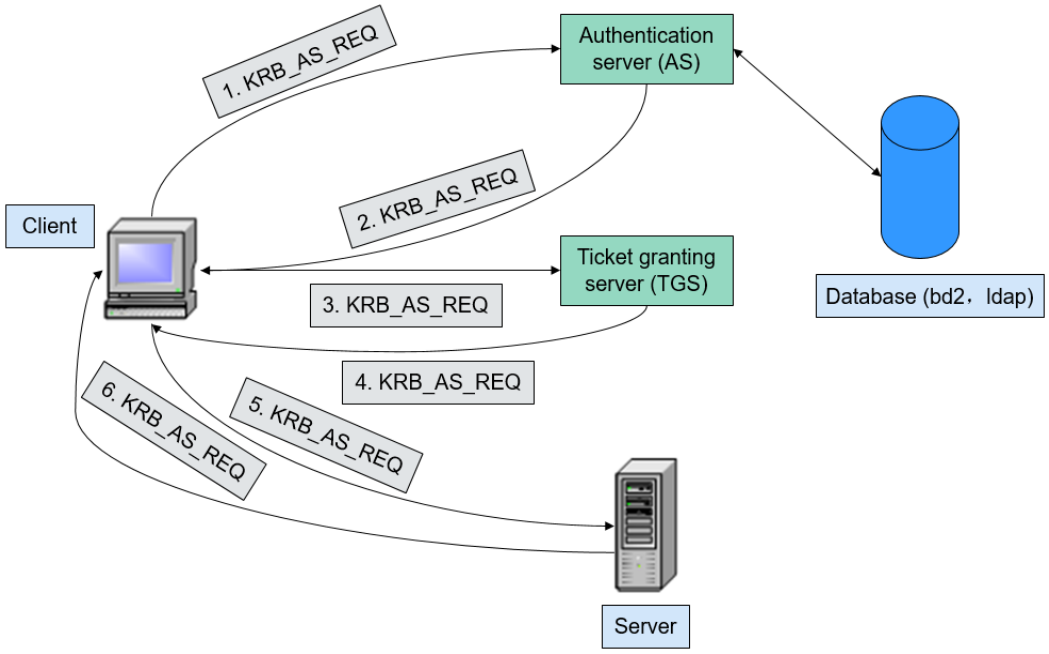
Kerberos authentication

Kerberos is a security authentication system designed with the client/server architecture that employs encryption technologies such as DES and AES. The system supports mutual authentication, that is, the client and server can authenticate identity of each other. Kerberos prevents interception and replay attacks, and protects data integrity. It manages keys by using a symmetric key mechanism.

Kerberos authentication consists of the following roles:

- Client
- Server
- Key Distribution Center (KDC): consists of the Authentication Server (AS) and Ticket Granting Server (TGS).
  - AS verifies the client account and password and generates a Ticket Granting Ticket (TGT).
  - TGS generates Service Tickets (STs) for accessing services based on TGTs.

Figure 1-78 Authentication process

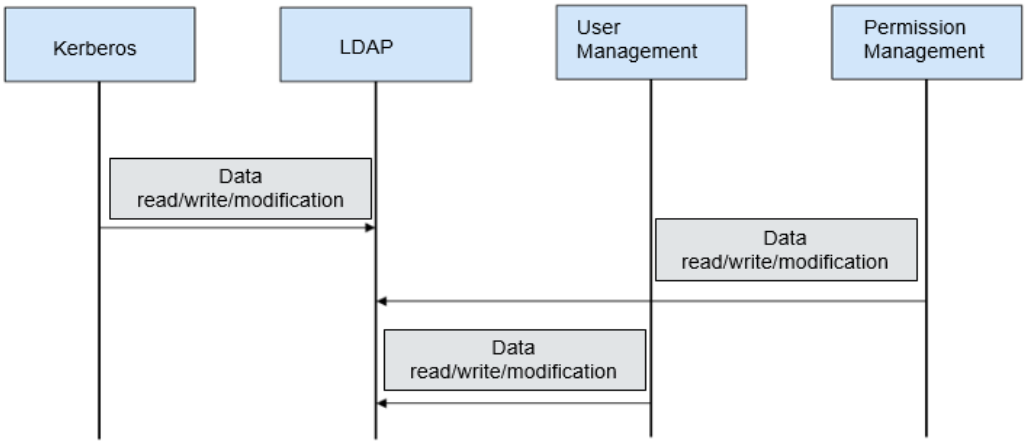


LDAP data read and write

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

LDAP serves as a user data storage center and stores user information (such as passwords and supplementary information) in the cluster. Users need to access LDAP to operate user data or perform Kerberos authentication.

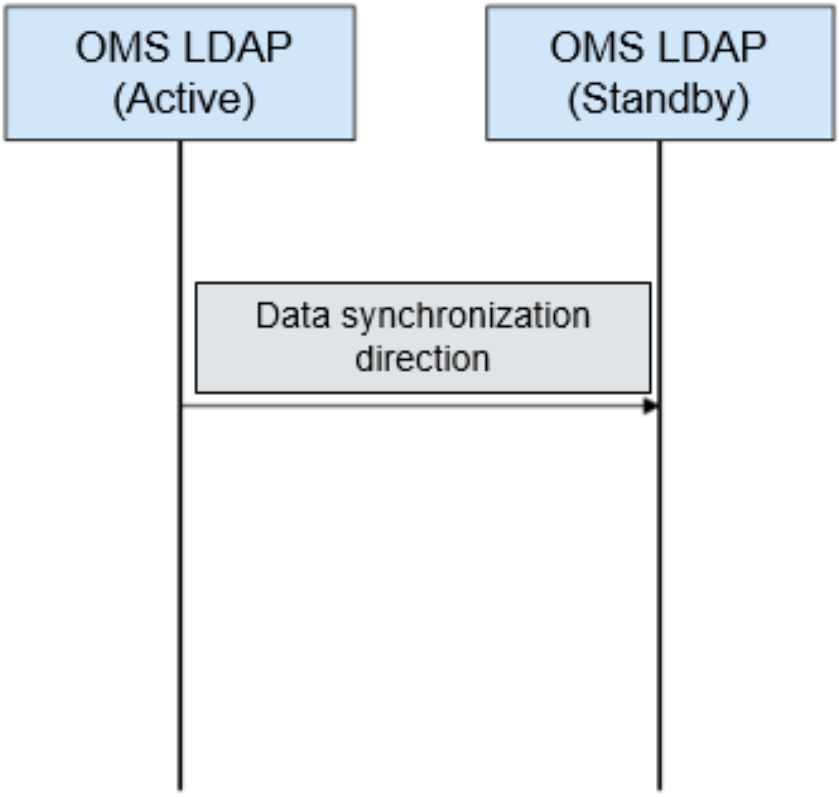
Figure 1-79 Data modification process



LDAP data synchronization

- OMS LDAP data synchronization before cluster installation

  Figure 1-80 OMS LDAP data synchronization

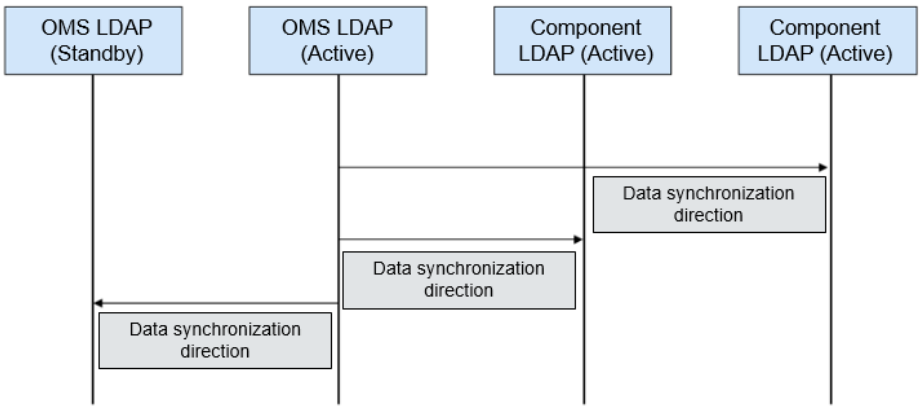


  Data synchronization direction before cluster installation: Data is synchronized from the active OMS LDAP to the standby OMS LDAP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- LDAP data synchronization after cluster installation

Figure 1-81 LDAP data synchronization



Data synchronization direction after cluster installation: Data is synchronized from the active OMS LDAP to the standby OMS LDAP, standby component LDAP, and standby component LDAP.

1.6.19.2 KrbServer and LdapServer Enhanced Open Source Features

Service Authentication in a Cluster

In an MRS cluster that uses the security mode, mutual access between services is implemented based on the Kerberos security architecture. When a service (such as HDFS) in the cluster is to be started, the corresponding sessionkey (keytab, used for identity authentication of the application) is obtained from Kerberos. If another service (such as YARN) needs to access HDFS and add, delete, modify, or query data in HDFS, the corresponding TGT and ST must be obtained for secure access.

Application Development Authentication

MRS components provide application development interfaces for customers or upper-layer service clusters. During application development, a cluster in security mode provides specified application development authentication interfaces to implement application security authentication and access. For example, the **UserGroupInformation** class provided in **hadoop-common api** includes multiple security authentication APIs.

- **setConfiguration()** is used to obtain related configuration and set parameters such as global variables.
- **loginUserFromKeytab():** is used to obtain TGT interfaces.

Cross-System Mutual Trust

MRS provides the mutual trust function between two Managers to implement data read and write operations between systems.

1.6.20 Kudu

Kudu is a column-store manager developed for the Apache Hadoop platform. Kudu shares the common technical properties of Hadoop ecosystem applications, that is, it runs on commodity hardware, which is horizontally scalable, delivering high availability.

Kudu's design has the following benefits:

- Fast processing of OLAP workloads
- Integration with MapReduce, Spark and other Hadoop ecosystem components
- Tight integration with Apache Impala, making it a good, mutable alternative to using HDFS with Apache Parquet
- Strong but flexible consistency model, allowing you to choose consistency requirements on a per-request basis, including the option for strict-serializable consistency
- Strong performance for running sequential and random workloads simultaneously
- Easy to manage
- High availability Tablet Servers and Masters use the Raft Consensus Algorithm, which ensures that as long as more than half the total number of replicas is available, the tablet is available for reads and writes. For example, if 2 out of 3 replicas or 3 out of 5 replicas are available, the tablet is available. Reads can be serviced by read-only follower tablets, even in the event of a leader tablet failure.
- Structured data model

By combining all of these properties, Kudu targets support for families of applications that are difficult or impossible to implement on current generation Hadoop storage technologies.

A few examples of applications for which Kudu is a great solution are:

- Reporting applications where newly-arrived data needs to be immediately available for end users
- Time-series applications that must simultaneously support queries across large amounts of historic data and granular queries about an individual entity that must return very quickly
- Applications that use predictive models to make real-time decisions with periodic refreshes of the predictive model based on all historic data

Relationship with Other Components

- HBase

  Kudu is designed based on the HBase structure and can implement fast random read/write and update functions that HBase features.

  The differences are as follows:

  - HBase uses ZooKeeper to ensure data consistency, whereas Kudu uses the Raft consensus algorithm to ensure consistency.
  - HBase uses HDFS for resilient data storage, whereas Kudu uses TServer to ensure strong data consistency and reliability.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.21 Loader

1.6.21.1 Loader Basic Principles

Loader is developed based on the open source Sqoop component. It is used to exchange data and files between MRS and relational databases and file systems. Loader can import data from relational databases or file servers to the HDFS and HBase components, or export data from HDFS and HBase to relational databases or file servers.

A Loader model consists of Loader Client and Loader Server, as shown in **Figure 1-82**.

Figure 1-82 Loader model

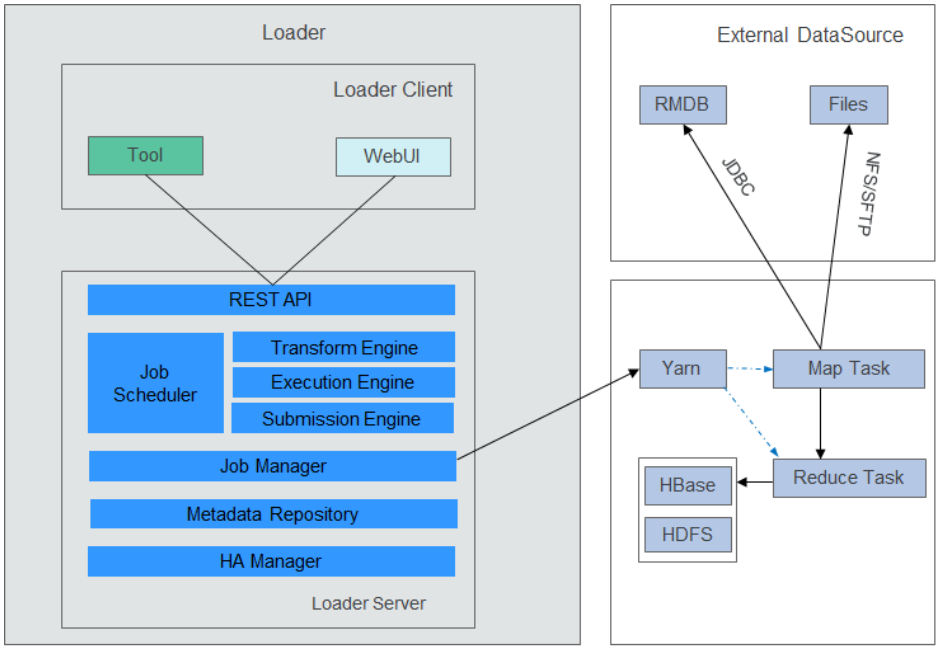


Table 1-19 describes the functions of each module shown in the preceding figure.

Table 1-19 Components of the Loader model

Module	Description
Loader Client	Loader client. It provides two interfaces: web UI and CLI.
Loader Server	Loader server. It processes operation requests sent from the client, manages connectors and metadata, submits MapReduce jobs, and monitors MapReduce job status.
REST API	It provides a Representational State Transfer (RESTful) APIs (HTTP + JSON) to process the operation requests sent from the client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
Job Scheduler	Simple job scheduler. It periodically executes Loader jobs.
Transform Engine	Data transformation engine. It supports field combination, string cutting, and string reverse.
Execution Engine	Loader job execution engine. It executes Loader jobs in MapReduce manner.
Submission Engine	Loader job submission engine. It submits Loader jobs to MapReduce.
Job Manager	It manages Loader jobs, including creating, querying, updating, deleting, activating, deactivating, starting, and stopping jobs.
Metadata Repository	Metadata repository. It stores and manages data about Loader connectors, transformation procedures, and jobs.
HA Manager	It manages the active/standby status of Loader Server processes. The Loader Server has two nodes that are deployed in active/standby mode.

Loader imports or exports jobs in parallel using MapReduce jobs. Some job import or export may involve only the Map operations, while some may involve both Map and Reduce operations.

Loader implements fault tolerance using MapReduce. Jobs can be rescheduled upon a job execution failure.

- **Importing data to HBase**

  When the Map operation is performed for MapReduce jobs, Loader obtains data from an external data source.

  When a Reduce operation is performed for a MapReduce job, Loader enables the same number of Reduce tasks based on the number of Regions. The Reduce tasks receive data from Map tasks, generate HFiles by Region, and store the HFiles in a temporary directory of HDFS.

  When a MapReduce job is submitted, Loader migrates HFiles from the temporary directory to the HBase directory.

- **Importing Data to HDFS**

  When a Map operation is performed for a MapReduce job, Loader obtains data from an external data source and exports the data to a temporary directory (named *export directory*-**ldtmp**).

  When a MapReduce job is submitted, Loader migrates data from the temporary directory to the output directory.

- **Exporting data to a relational database**

  When a Map operation is performed for a MapReduce job, Loader obtains data from HDFS or HBase and inserts the data to a temporary table (Staging Table) through the Java DataBase Connectivity (JDBC) API.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

When a MapReduce job is submitted, Loader migrates data from the temporary table to a formal table.

- **Exporting data to a file system**

  When a Map operation is performed for a MapReduce job, Loader obtains data from HDFS or HBase and writes the data to a temporary directory of the file server.

  When a MapReduce job is submitted, Loader migrates data from the temporary directory to a formal directory.

1.6.21.2 Relationship Between Loader and Other Components

The components that interact with Loader include HDFS, HBase, Hive, YARN, MapReduce, and ZooKeeper.

Loader works as a client and uses certain functions of these components, such as storing data to HDFS and HBase and reading data from HDFS and HBase tables. In addition, Loader functions as a MapReduce client to import or export data.

Loader imports or exports jobs in parallel using MapReduce jobs. Some job import or export may involve only the Map operations, while some may involve both Map and Reduce operations.

1.6.21.3 Loader Enhanced Open Source Features

Loader Enhanced Open-Source Feature: Data Import and Export

Loader is developed based on Sqoop. In addition to the Sqoop functions, Loader has the following enhanced features:

- Provides data conversion functions.
- Supports GUI-based configuration conversion.
- Imports data from an SFTP/FTP server to HDFS/OBS.
- Imports data from an SFTP/FTP server to an HBase table.
- Imports data from an SFTP/FTP server to a Phoenix table.
- Imports data from an SFTP/FTP server to a Hive table.
- Exports data from HDFS/OBS to an SFTP server.
- Exports data from an HBase table to an SFTP server.
- Exports data from a Phoenix table to an SFTP server.
- Imports data from a relational database to an HBase table.
- Imports data from a relational database to a Phoenix table.
- Imports data from a relational database to a Hive table.
- Exports data from an HBase table to a relational database.
- Exports data from a Phoenix table to a relational database.
- Imports data from an Oracle partitioned table to HDFS/OBS.
- Imports data from an Oracle partitioned table to an HBase table.
- Imports data from an Oracle partitioned table to a Phoenix table.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Imports data from an Oracle partitioned table to a Hive table.
- Exports data from HDFS/OBS to an Oracle partitioned table.
- Exports data from HBase to an Oracle partitioned table.
- Exports data from a Phoenix table to an Oracle partitioned table.
- Imports data from HDFS to an HBase table, a Phoenix table, and a Hive table in the same cluster.
- Exports data from an HBase table and a Phoenix table to HDFS/OBS in the same cluster.
- Imports data to an HBase table and a Phoenix table by using **bulkload** or **put list**.
- Imports all types of files from an SFTP/FTP server to HDFS. The open source component Sqoop can import only text files.
- Exports all types of files from HDFS/OBS to an SFTP server. The open source component Sqoop can export only text files and SequenceFile files.
- Supports file coding format conversion during file import and export. The supported coding formats include all formats supported by Java Development Kit (JDK).
- Retains the original directory structure and file names during file import and export.
- Supports file combination during file import and export. For example, if a large number of files are to be imported, these files can be combined into *n* files (*n* can be configured).
- Supports file filtering during file import and export. The filtering rules support wildcards and regular expressions.
- Supports batch import and export of ETL tasks.
- Supports query by page and key word and group management of ETL tasks.
- Provides floating IP addresses for external components.

1.6.22 Manager

1.6.22.1 Manager Basic Principles

Overview

Manager is the O&M management system of MRS and provides unified cluster management capabilities for services deployed in clusters.

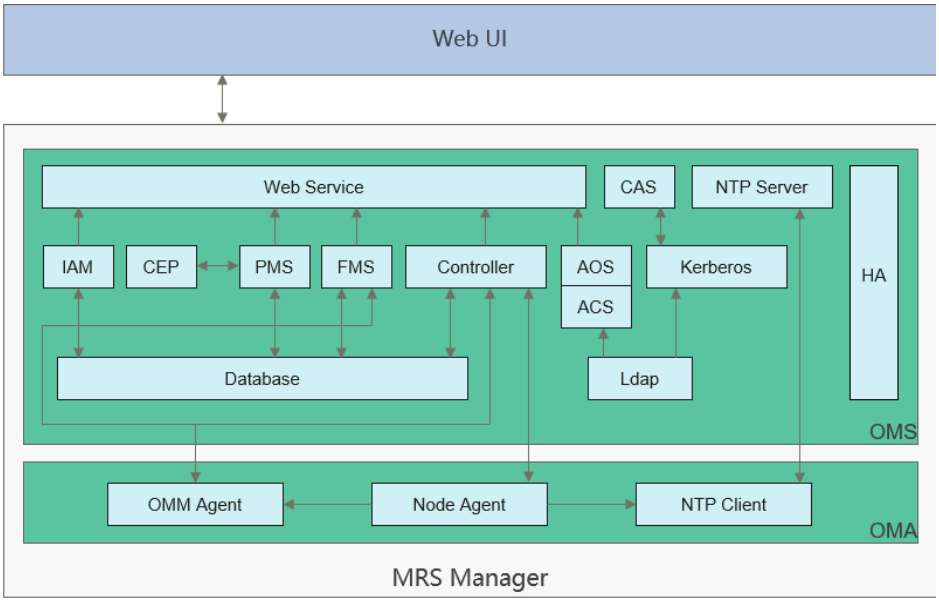
Manager provides functions such as performance monitoring, alarms, user management, permission management, auditing, service management, health check, and log collection.

Architecture

Figure 1-83 shows the overall logical architecture of FusionInsight Manager.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-83 Manager logical architecture



Manager consists of OMS and OMA.

- OMS: serves as management node in the O&M system. There are two OMS nodes deployed in active/standby mode.
- OMA: managed node in the O&M system. Generally, there are multiple OMA nodes.

Table 1-20 describes the modules shown in **Figure 1-83**.

Table 1-20 Service module description

Module	Description
Web Service	A web service deployed under Tomcat, providing HTTPS API of Manager. It is used to access Manager through the web browser. In addition, it provides the northbound access capability based on the Syslog and SNMP protocols.
OMS	Management node of the O&M system. Generally, there are two OMS nodes that work in active/standby mode.
OMA	Managed node in the O&M system. Generally, there are multiple OMA nodes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
Controller	The control center of Manager. It can converge information from all nodes in the cluster and display it to MRS cluster administrators, as well as receive from MRS cluster administrators, and synchronize information to all nodes in the cluster according to the operation instruction range. Control process of Manager. It implements various management actions: 1. The web service delivers various management actions (such as installation, service startup and stop, and configuration modification) to Controller. 2. Controller decomposes the command and delivers the action to each Node Agent, for example, starting a service involves multiple roles and instances. 3. Controller is responsible for monitoring the implementation of each action.
Node Agent	Node Agent exists on each cluster node and is an enabler of Manager on a single node. • Node Agent represents all the components deployed on the node to interact with Controller, implementing convergence from multiple nodes of a cluster to a single node. • Node Agent enables Controller to perform all operations on the components deployed on the node. It allows Controller functions to be implemented. Node Agent sends heartbeat messages to Controller at an interval of 3 seconds. The interval cannot be configured.
IAM	Records audit logs. Each non-query operation on the Manager UI has a related audit log.
PMS	The performance monitoring module. It collects the performance monitoring data on each OMA and provides the query function.
CEP	Convergence function module. For example, the used disk space of all OMAs is collected as a performance indicator.
FMS	Alarm module. It collects and queries alarms on each OMA.
OMM Agent	Agent for performance monitoring and alarm reporting on the OMA. It collects performance monitoring data and alarm data on Agent Node.
CAS	Unified authentication center. When a user logs in to the web service, CAS authenticates the login. The browser automatically redirects the user to the CAS through URLs.
AOS	Permission management module. It manages the permissions of users and user groups.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
ACS	User and user group management module. It manages users and user groups to which users belong.
Kerberos	LDAP is deployed in OMS and a cluster, respectively. • OMS Kerberos provides the single sign-on (SSO) and authentication between Controller and Node Agent. • Kerberos in the cluster provides the user security authentication function for components. The service name is **KrbServer**, which contains two role instances: – KerberosServer: is an authentication server that provides security authentication for MRS. – KerberosAdmin: manages processes of Kerberos users.
Ldap	LDAP is deployed in OMS and a cluster, respectively. • OMS LDAP provides data storage for user authentication. • The LDAP in the cluster functions as the backup of the OMS LDAP. The service name is **LdapServer** and the role instance is **SlapdServer**.
Database	Manager database used to store logs and alarms.
HA	HA management module that manages the active and standby OMSs.
NTP Server NTP Client	It synchronizes the system clock of each node in the cluster.

1.6.22.2 Manager Key Features

Key Feature: Unified Alarm Monitoring

Manager provides the visualized and convenient alarm monitoring function. Users can quickly obtain key cluster performance indicators, evaluate cluster health status, customize performance indicator display, and convert indicators to alarms. Manager can monitor the running status of all components and report alarms in real time when faults occur. The online help on the GUI allows you to view performance counters and alarm clearance methods to quickly rectify faults.

Key Feature: Unified User Permission Management

Manager provides permission management of components in a unified manner.

Manager introduces the concept of role and uses role-based access control (RBAC) to manage system permissions. It centrally displays and manages scattered permission functions of each component in the system and organizes the permissions of each component in the form of permission sets (roles) to form a unified system permission concept. By doing so, common users cannot obtain internal permission management details, and permissions become easy for MRS

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cluster administrators to manage, greatly facilitating permission management and improving user experience.

Key Feature: SSO

Single sign-on (SSO) is provided between the Manager web UI and component web UI as well as for integration between MRS and third-party systems.

This function centrally manages and authenticates Manager users and component users. The entire system uses LDAP to manage users and uses Kerberos for authentication. A set of Kerberos and LDAP management mechanisms are used between the OMS and components. SSO (including single sign-on and single sign-out) is implemented through CAS. With SSO, users can easily switch tasks between the Manager web UI, component web UIs, and third-party systems, without switching to another user.

NOTE

- To ensure security, the CAS Server can retain a ticket-granting ticket (TGT) used by a user only for 20 minutes.
- If a user does not perform any operation on the page (including on the Manager web UI and component web UIs) within 20 minutes, the page is automatically locked.

Key Feature: Automatic Health Check and Inspection

Manager provides users with automatic inspection on system running environments and helps users check and audit system running health by one click, ensuring correct system running and lowering system operation and maintenance costs. After viewing inspection results, you can export reports for archiving and fault analysis.

Key Feature: Tenant Management

Manager introduces the multi-tenant concept. The CPU, memory, and disk resources of a cluster can be integrated into a set. The set is called a tenant. A mode involving different tenants is called multi-tenant mode.

Manager provides the multi-tenant function, supports a level-based tenant model and allows tenants to be added and deleted dynamically, achieving resource isolation. As a result, it can dynamically manage and configure the computing resources and the storage resources of tenants.

- The computing resources indicate tenants' Yarn task queue resources. The task queue quota can be modified, and the task queue usage status and statistics can be viewed.
- The storage resources can be stored on HDFS. You can add and delete the HDFS storage directories of tenants, and set the quotas of file quantity and the storage space of the directories.

As a unified tenant management platform of MRS, MRS Manager allows users to create and manage tenants in clusters based on service requirements.

- Roles, computing resources, and storage resources are automatically created when tenants are created. By default, all permissions of the new computing resources and storage resources are allocated to a tenant's roles.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After you have modified the tenant's computing or storage resources, permissions of the tenant's roles are automatically updated.

Manager also provides the multi-instance function so that users can use the HBase, Hive, or Spark alone in the resource control and service isolation scenario. The multi-instance function is disabled by default and can be manually enabled.

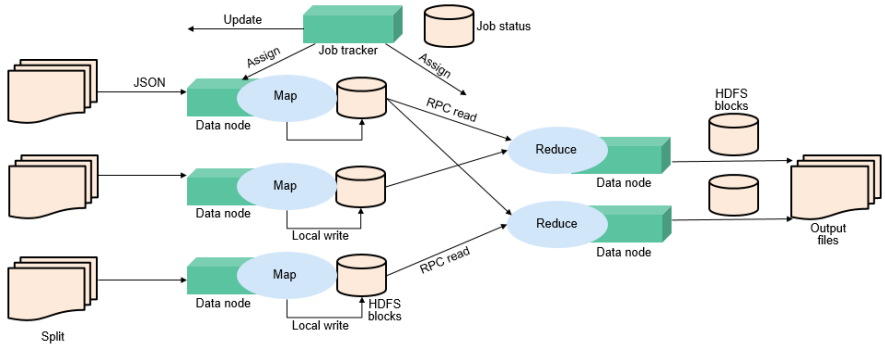
1.6.23 MapReduce

1.6.23.1 MapReduce Basic Principles

MapReduce is the core of Hadoop. As a software architecture proposed by Google, MapReduce is used for parallel computing of large-scale datasets (larger than 1 TB). The concepts "Map" and "Reduce" and their main ideas are derived from functional programming languages and vector programming languages.

Current software implementation is as follows: Specify a Map function to map a series of key-value pairs into a new series of key-value pairs, and specify a Reduce function to ensure that all values in the mapped key-value pairs share the same key.

Figure 1-84 Distributed batch processing engine



MapReduce is a software framework for processing large datasets in parallel. The root of MapReduce is the Map and Reduce functions in functional programming. The Map function accepts a group of data and transforms it into a key-value pair list. Each element in the input domain corresponds to a key-value pair. The Reduce function accepts the list generated by the Map function, and then shrinks the key-value pair list based on the keys. MapReduce divides a task into multiple parts and allocates them to different devices for processing. In this way, the task can be finished in a distributed environment instead of a single powerful server.

NOTE

To use MapReduce, ensure that the Hadoop service has been installed in the MRS cluster.

MapReduce structure

MapReduce is integrated into YARN through the Client and ApplicationMaster interfaces of YARN, and uses YARN to apply for computing resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.23.2 Relationship Between MapReduce and Other Components

Relationship Between MapReduce and HDFS

- HDFS features high fault tolerance and high throughput, and can be deployed on low-cost hardware for storing data of applications with massive data sets.
- MapReduce is a programming model used for parallel computation of large data sets (larger than 1 TB). Data computed by MapReduce comes from multiple data sources, such as Local FileSystem, HDFS, and databases. HDFS is most commonly used. MapReduce uses the high throughput performance of HDFS to read a large amount of data for computing and the data is stored in HDFS.

Relationship Between MapReduce and Yarn

MapReduce is a computing framework running on Yarn, which is used for batch processing. MRv1 is implemented based on MapReduce in Hadoop 1.0, which is composed of programming models (new and old programming APIs), running environment (JobTracker and TaskTracker), and data processing engine (MapTask and ReduceTask). This framework is still weak in scalability, fault tolerance (JobTracker SPOF), and compatibility with multiple frameworks. (Currently, only the MapReduce computing framework is supported.) MRv2 is implemented based on MapReduce in Hadoop 2.0. The source code reuses MRv1 programming models and data processing engine implementation, and the running environment is composed of ResourceManager and ApplicationMaster. ResourceManager is a brand new resource manager system, and ApplicationMaster is responsible for cutting MapReduce job data, assigning tasks, applying for resources, scheduling tasks, and tolerating faults.

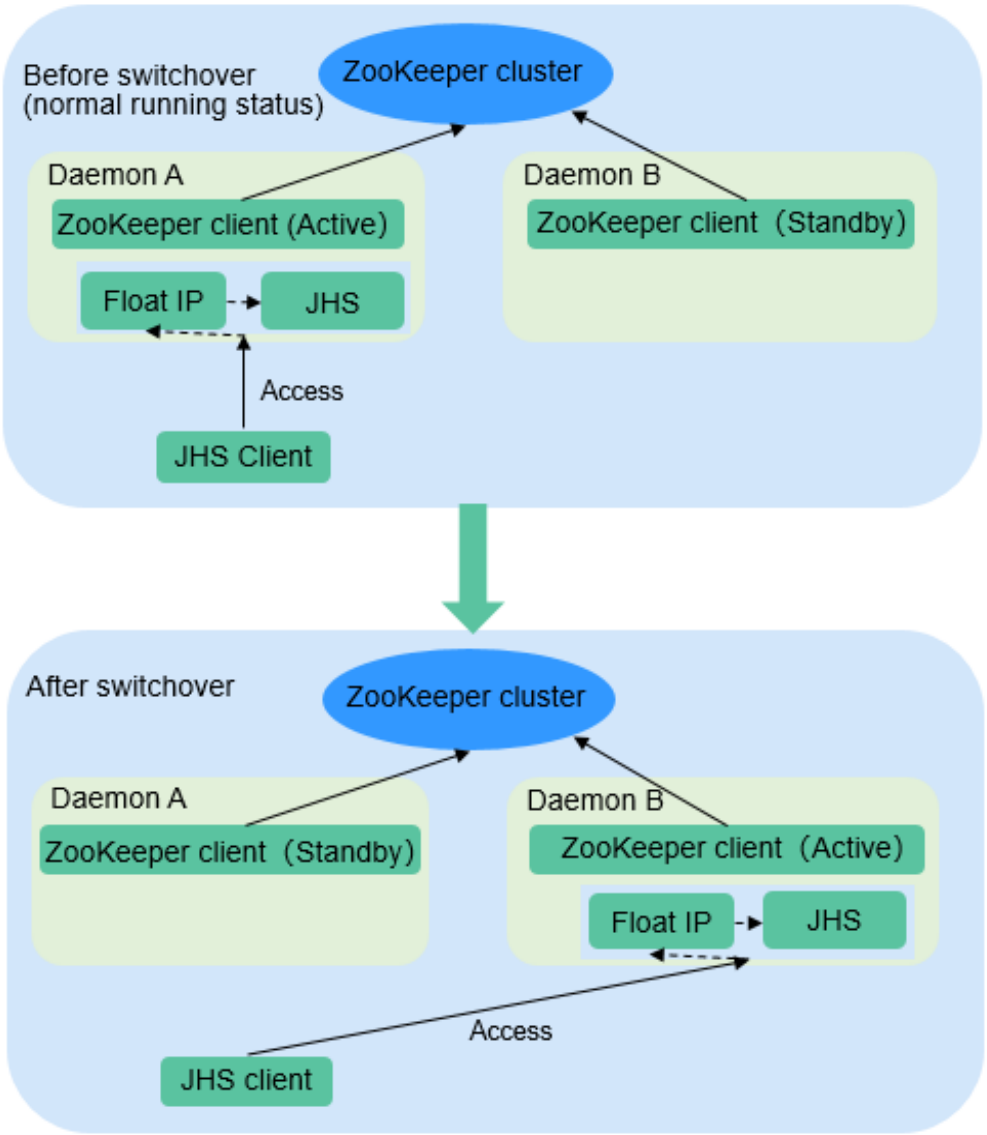
1.6.23.3 MapReduce Enhanced Open Source Features

MapReduce Enhanced Open-Source Feature: JobHistoryServer HA

JobHistoryServer (JHS) is the server used to view historical MapReduce task information. Currently, the open source JHS supports only single-instance services. JHS HA can solve the problem that an application fails to access the MapReduce API when SPOFs occur on the JHS, which causes the application fails to be executed. This greatly improves the high availability of the MapReduce service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-85 Status transition of the JobHistoryServer HA active/standby switchover



JobHistoryServer High Availability

- ZooKeeper is used to implement active/standby election and switchover.
- JHS uses the floating IP address to provide services externally.
- Both the JHS single-instance and HA deployment modes are supported.
- Only one node starts the JHS process at a time point to prevent multiple JHS operations from processing the same file.
- You can perform scale-out, scale-in, instance migration, upgrade, and health check.

Enhanced Open Source Feature: Improving MapReduce Performance by Optimizing the Merge/Sort Process in Specific Scenarios

The figure below shows the workflow of a MapReduce task.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-86 MapReduce job

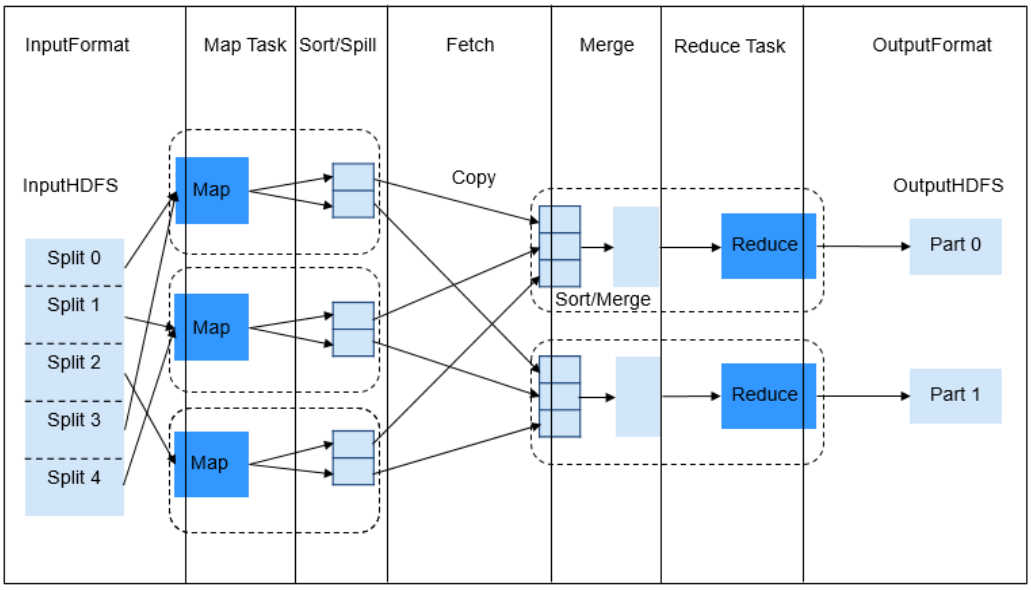
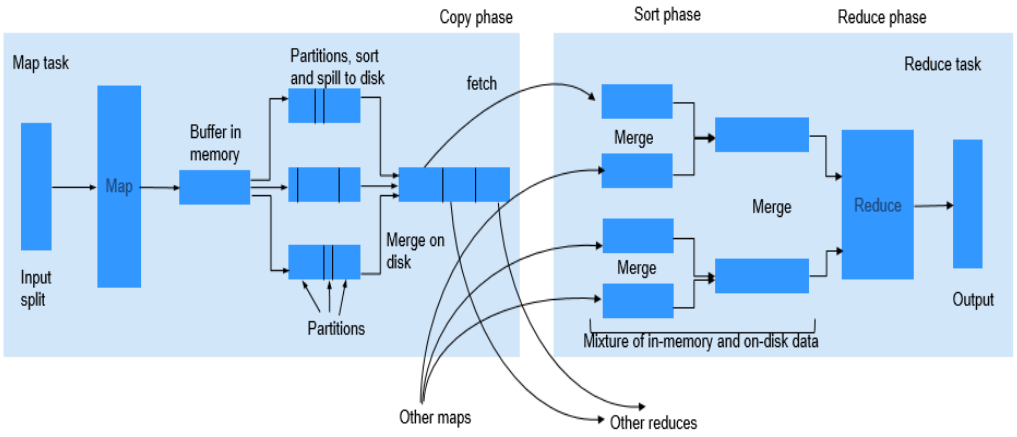


Figure 1-87 Job execution process



The Reduce process is divided into three different steps: Copy, Sort (actually supposed to be called Merge), and Reduce. In Copy phase, Reducer tries to fetch the output of Maps from NodeManagers and store it on Reducer either in memory or on disk. Shuffle (Sort and Reduce) phase then begins. All the fetched maps are sorted, and segments from different maps are merged before being sent to Reducer. When a job has a large number of maps to be processed, the shuffle process is time-consuming. For specific tasks (for example, SQL tasks such as hash join and hash aggregation), sorting is not mandatory during the shuffle process. However, the sorting is required by default in the shuffle process.

This feature is enhanced by using the MapReduce API, which can automatically close the Sort process for such tasks. When the sorting is disabled, the API directly merges the fetched Maps output data and sends the data to Reducer. This greatly saves time, and significantly improves the efficiency of SQL tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Enhanced Open Source Feature: Small Log File Problem Solved After Optimization of History Server

After the job running on Yarn is executed, NodeManager uses LogAggregationService to collect and send generated logs to HDFS and deletes them from the local file system. After the logs are stored to HDFS, they are managed by HistoryServer. LogAggregationService will merge local logs generated by containers to a log file and upload it to the HDFS, reducing the number of log files to some extent. However, in a large-scale and busy cluster, there will be excessive log files on HDFS after long-term running.

For example, if there are 20 nodes, about 18 million log files are generated within the default clean-up period (15 days), which occupy about 18 GB of the memory of a NameNode and slow down the HDFS system response.

Only the reading and deletion are required for files stored on HDFS. Therefore, Hadoop Archives can be used to periodically archive the directory of collected log files.

Archiving Logs

The AggregatedLogArchiveService module is added to HistoryServer to periodically check the number of files in the log directory. When the number of files reaches the threshold, AggregatedLogArchiveService starts an archiving task to archive log files. After archiving, it deletes the original log files to reduce log files on HDFS.

Cleaning Archived Logs

Hadoop Archives does not support deletion in archived files. Therefore, the entire archive log package must be deleted upon log clean-up. The latest log generation time is obtained by modifying the AggregatedLogDeletionService module. If all log files meet the clean-up requirements, the archive log package can be deleted.

Browsing Archived Logs

Hadoop Archives allows URI-based access to file content in the archive log package. Therefore, if History Server detects that the original log files do not exist during file browsing, it directly redirects the URI to the archive log package to access the archived log file.

NOTE

- This function invokes Hadoop Archives of HDFS for log archiving. Because the execution of an archiving task by Hadoop Archives is to run an MR application. Therefore, after an archiving task is executed, an MR execution record is added.
- This function of archiving logs is based on the log collection function. Therefore, this function is valid only when the log collection function is enabled.

1.6.24 MemArtsCC

1.6.24.1 MemArtsCC Basic Principles

MemArtsCC is a distributed lightweight caching system deployed on the compute side in the storage-compute decoupled architecture. It prefetches data from remote object storage to accelerate caching and computing jobs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MemArtsCC shards objects on remote OBS and creates indexes, greatly improving the performance of reading cached data. ZooKeeper enables lightweight service discovery and provides ultra-high availability. The lifecycle management of sharded data is based on the LRU algorithm.

Main Features of MemArtsCC

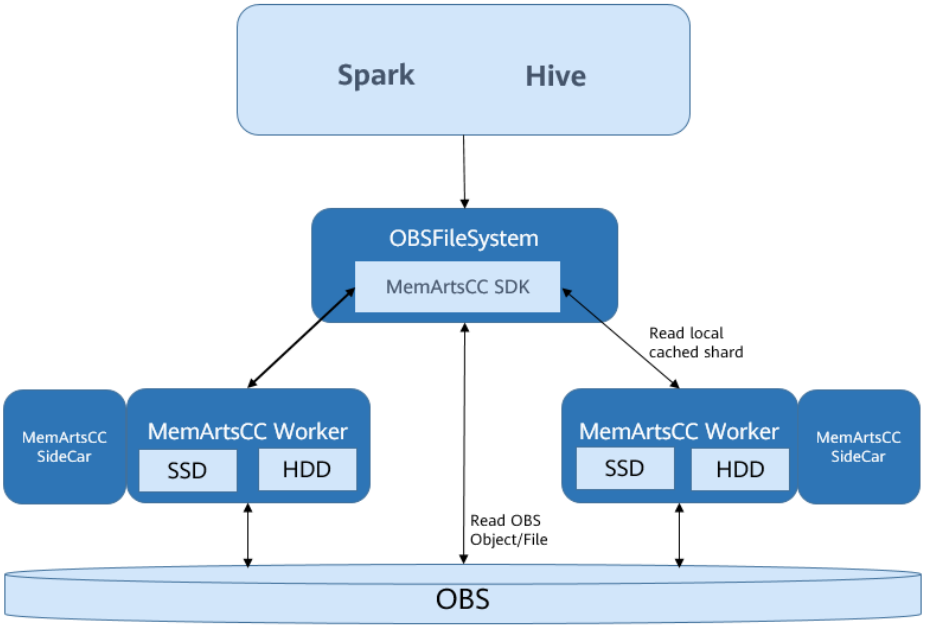
- The decentralized architecture enables all instances to provide same service capabilities.
- The lightweight design ensures ultra-low resource usage.
- MemArtsCC is decoupled from applications and therefore is transparent to them and can be used without adaptation.
- MemArtsCC ensures high availability in case of node failures.

MemArtsCC Structure

There are CCSideCar and CCWorker roles of MemArtsCC instances.

In a storage-compute decoupled architecture, data of computing and analytics applications such as Spark and Hive is stored in OBS. In a MemArtsCC cluster, a service instance is called a worker. Workers cache some or all of the object data in OBS to local persistent storage (SSD/HDD). When an application reads an object through the MemArtsCC SDK, the application reads sharded data from a specific worker based on the shard index. If the cache is hit, the worker returns the shards. If the cache is not hit, the application directly reads data from OBS. The worker asynchronously loads the shards that are not hit to local storage for subsequent use.

Figure 1-88 MemArtsCC Structure



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-21 MemArtsCC architecture

Name	Description
MemArtsCC SDK	An SDK used by OBSA (a Hadoop client plug-in) to access OBS server objects.
CCSideCar	The management plane service. It monitors MemArtsCC, collects data, delivers configurations, and starts and stops the service.
CCWorker	The data plane service. It reads, writes, stores, and deletes data cached by MemArtsCC.

1.6.24.2 Relationships Between MemArtsCC and Other Components

OBS

OBS provides a new InputStream: OBSMemArtsCCInputStream. This InputStream reads data from the MemArtsCC cluster deployed on the compute side to reduce OBS server pressure and improve data read performance.

MemArtsCC persistently stores data to SSDs on the compute side. OBS interconnects with MemArtsCC to:

1. Improve data access performance in the decoupled storage-compute architecture.

   The local storage of MemArtsCC avoids the cross-network access to hotspot data. This accelerates the data reads of OBS upper-layer applications.

2. Reduce the pressure on the OBS server.

   MemArtsCC stores hotspot data in the compute cluster to reduce the bandwidth pressure of the OBS server.

Spark

Spark reads data from OBS. OBS reads data from MemArtsCC. If data is hit in the local cache, the data is read directly. Otherwise, the data is prefetched.

Hive

Hive reads data from OBS. OBS reads data from MemArtsCC. If data is hit in the local cache, the data is read directly. Otherwise, the data is prefetched.

HetuEngine

HetuEngine reads data from OBS. OBS reads data from MemArtsCC. If data is hit in the local cache, the data is read directly. Otherwise, the data is prefetched.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.25 Oozie

Introduction to Oozie

Oozie is an open-source workflow engine that is used to schedule and coordinate Hadoop jobs.

Architecture

The Oozie engine is a web application integrated into Tomcat by default. Oozie uses PostgreSQL databases.

Oozie provides an Ext-based web console, through which users can view and monitor Oozie workflows. Oozie provides an external REST web service API for the Oozie client to control workflows (such as starting and stopping operations), and orchestrate and run Hadoop MapReduce tasks. For details, see **Figure 1-89**.

Figure 1-89 Oozie architecture

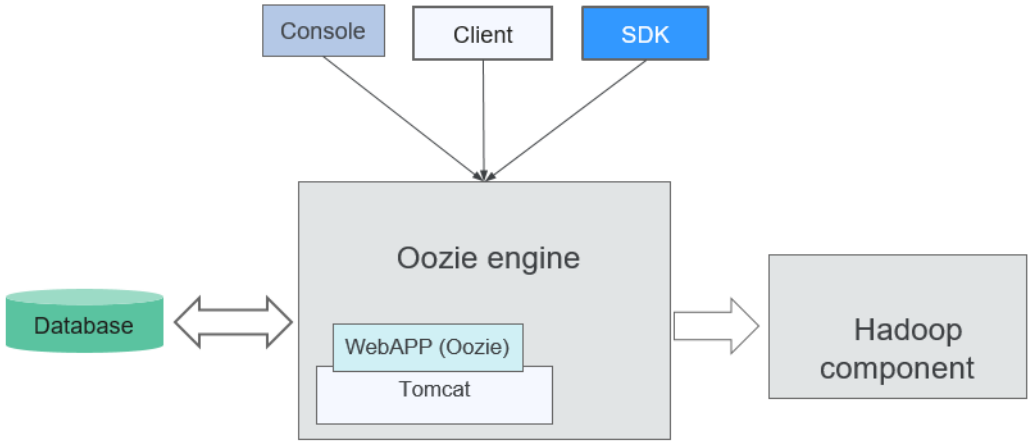


Table 1-22 describes the functions of each module shown in **Figure 1-89**.

Table 1-22 Architecture description

Connection Name	Description
Console	Allows users to view and monitor Oozie workflows.
Client	Controls workflows, including submitting, starting, running, planting, and restoring workflows, through APIs.
SDK	Is short for software development kit. An SDK is a set of development tools used by software engineers to establish applications for particular software packages, software frameworks, hardware platforms, and operating systems.
Database	PostgreSQL database

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Connection Name	Description
WebApp (Oozie)	Functions as the Oozie server. It can be deployed on a built-in or an external Tomcat container. Information recorded by WebApp (Oozie) including logs is stored in the PostgreSQL database.
Tomcat	A free open-source web application server
Hadoop components	Underlying components, such as MapReduce and Hive, that execute the workflows orchestrated by Oozie.

Principle

Oozie is a workflow engine server that runs MapReduce workflows. It is also a Java web application running in a Tomcat container.

Oozie workflows are constructed using Hadoop Process Definition Language (HPDL). HPDL is an XML-defined language, similar to JBoss jBPM Process Definition Language (jPDL). An Oozie workflow consists of the Control Node and Action Node.

- Control Node controls workflow orchestration, such as **start**, **end**, **error**, **decision**, **fork**, and **join**.
- An Oozie workflow contains multiple Action Nodes, such as MapReduce and Java.

  All Action Nodes are deployed and run in Direct Acyclic Graph (DAG) mode. Therefore, Action Nodes run in direction. That is, the next Action Node can run only when the running of the previous Action Node ends. When one Action Node ends, the remote server calls back the Oozie interface. Then Oozie executes the next Action Node of workflow in the same manner until all Action Nodes are executed (execution failures are counted).

Oozie workflows provide various types of Action Nodes, such as MapReduce, Hadoop distributed file system (HDFS), Secure Shell (SSH), Java, and Oozie sub-flows, to support a wide range of business requirements.

Enhanced Open Source Feature

Improved Security:

Provides roles of administrator and common users to support Oozie permission management.

Supports single sign-on and sign-out, HTTPS access, and audit logs.

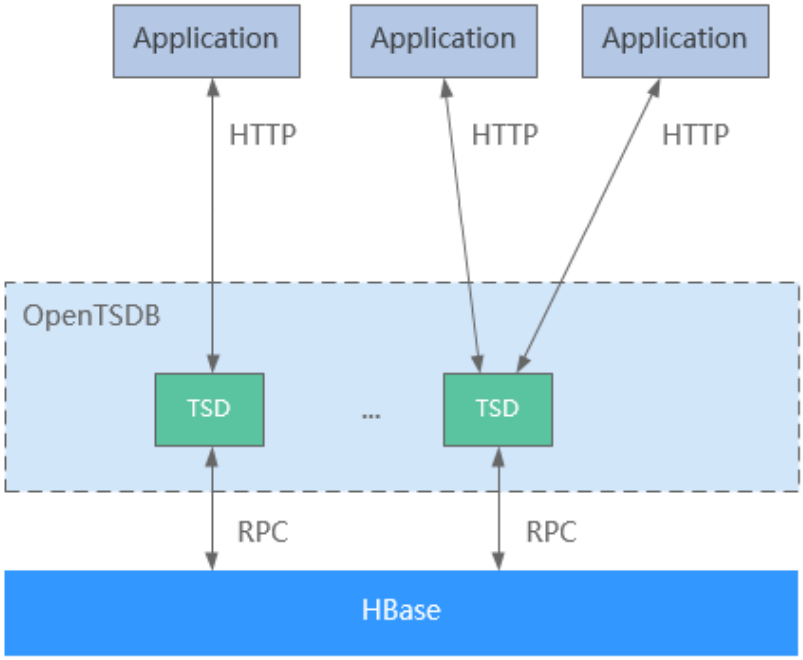
1.6.26 OpenTSDB

OpenTSDB is a distributed, scalable time series database based on HBase. OpenTSDB is designed to collect monitoring information of a large-scale cluster and implement second-level data query, eliminating the limitations of querying and storing massive amounts of monitoring data in common databases.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

OpenTSDB consists of a Time Series Daemon (TSD) as well as a set of command line utilities. Interaction with OpenTSDB is primarily implemented by running one or more TSDs. Each TSD is independent. There is no master server and no shared state, so you can run as many TSDs as required to handle any load you throw at it. Each TSD uses HBase in a CloudTable cluster to store and retrieve time series data. The data schema is highly optimized for fast aggregations of similar time series to minimize storage space. TSD users never need to directly access the underlying storage. You can communicate with the TSD through an HTTP API. All communications happen on the same port (the TSD figures out the protocol of the client by looking at the first few bytes it receives).

Figure 1-90 OpenTSDB architecture



Application scenarios of OpenTSDB have the following features:

- The collected metrics have a unique value at a time point and do not have a complex structure or relationship.
- Monitoring metrics change with time.
- Like HBase, OpenTSDB features high throughput and good scalability.

OpenTSDB provides an HTTP based application programming interface to enable integration with external systems. Almost all OpenTSDB features are accessible via the API such as querying time series data, managing metadata, and storing data points.

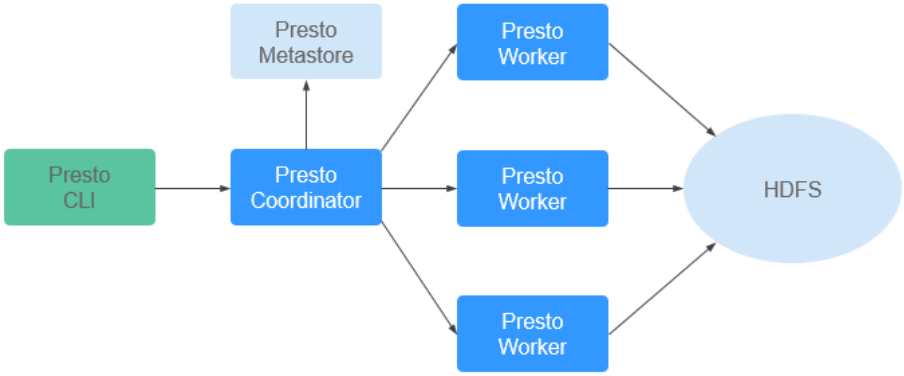
1.6.27 Presto

Presto is an open source SQL query engine for running interactive analytic queries against data sources of all sizes. It applies to massive structured/semi-structured data analysis, massive multi-dimensional data aggregation/report, ETL, ad-hoc queries, and more scenarios.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Presto allows querying data where it lives, including HDFS, Hive, HBase, Cassandra, relational databases or even proprietary data stores. A Presto query can combine different data sources to perform data analysis across the data sources.

Figure 1-91 Presto architecture



Presto runs in a cluster in distributed mode and contains one coordinator and multiple worker processes. Query requests are submitted from clients (for example, CLI) to the coordinator. The coordinator parses SQL statements, generates execution plans, and distributes the plans to multiple worker processes for execution.

Multiple Presto Instances

MRS supports multiple Presto instances. That is, multiple Worker instances, such as Worker1, Worker2, and Worker3, are installed on a Core/Task node. Multiple Worker instances interact with the Coordinator to execute computing tasks, greatly improving node resource utilization and computing efficiency.

Presto multi-instance applies only to the Arm architecture. Currently, a single node supports a maximum of four instances.

1.6.28 Ranger

1.6.28.1 Ranger Basic Principles

Apache Ranger offers a centralized security management framework and supports unified authorization and auditing. It manages fine grained access control over Hadoop and related components, such as Storm, HDFS, Hive, HBase, and Kafka. You can use the front-end web UI console provided by Ranger to configure policies to control users' access to these components.

Figure 1-92 shows the Ranger architecture.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-92 Ranger structure

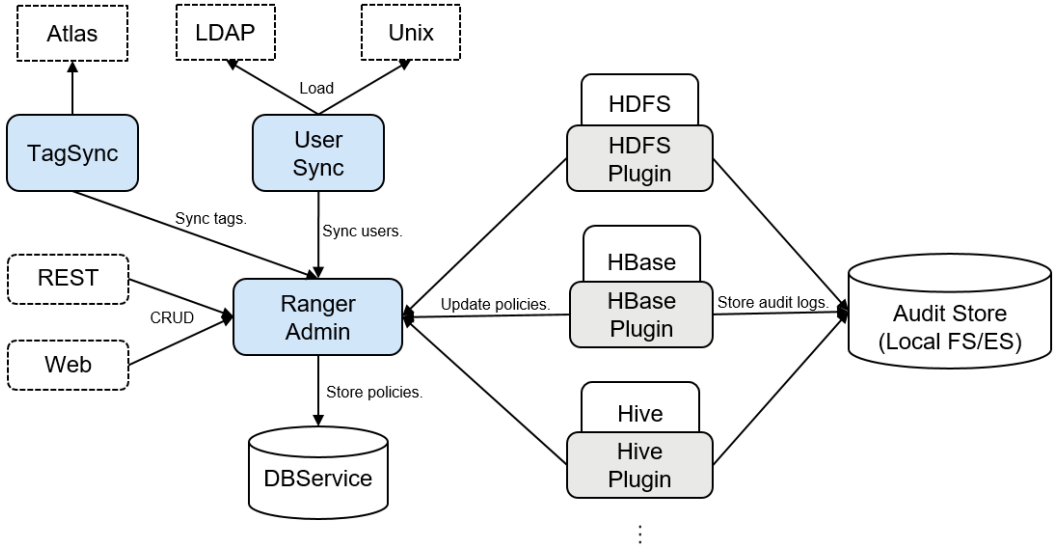


Table 1-23 Architecture description

Connection Name	Description
RangerAdmin	Provides a WebUI and RESTful API to manage policies, users, and auditing.
UserSync	Periodically synchronizes user and user group information from an external system and writes the information to RangerAdmin.
TagSync	Periodically synchronizes tag information from the external Atlas service and writes the tag information to RangerAdmin.

Ranger Principles

- Ranger Plugins

  Ranger provides policy-based access control (PBAC) plug-ins to replace the original authentication plug-ins of the components. Ranger plug-ins are developed based on the authentication interface of the components. Users set permission policies for specified services on the Ranger web UI. Ranger plug-ins periodically update policies from the RangerAdmin and caches them in the local file of the component. When a client request needs to be authenticated, the Ranger plug-in matches the user carried in the request with the policy and then returns an accept or reject message.

- UserSync User Synchronization

  UserSync periodically synchronizes data from LDAP/Unix to RangerAdmin. In security mode, data is synchronized from LDAP. In non-security mode, data is synchronized from Unix. By default, the incremental synchronization mode is used. In each synchronization period, UserSync updates only new or modified users and user groups. When a user or user group is deleted, UserSync does not synchronize the change to RangerAdmin. That is, the user or user group is

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

not deleted from the RangerAdmin. To improve performance, UserSync does not synchronize user groups to which no user belongs to RangerAdmin.

- Unified auditing

  Ranger plug-ins can record audit logs. Currently, audit logs can be stored in local files.

- High reliability

  Ranger supports two RangerAdmins working in active/active mode. Two RangerAdmins provide services at the same time. If either RangerAdmin is faulty, Ranger continues to work.

- High performance

  Ranger provides the Load-Balance capability. When a user accesses Ranger WebUI using a browser, the Load-Balance automatically selects the RangerAdmin with the lightest load to provide services.

1.6.28.2 Relationship Between Ranger and Other Components

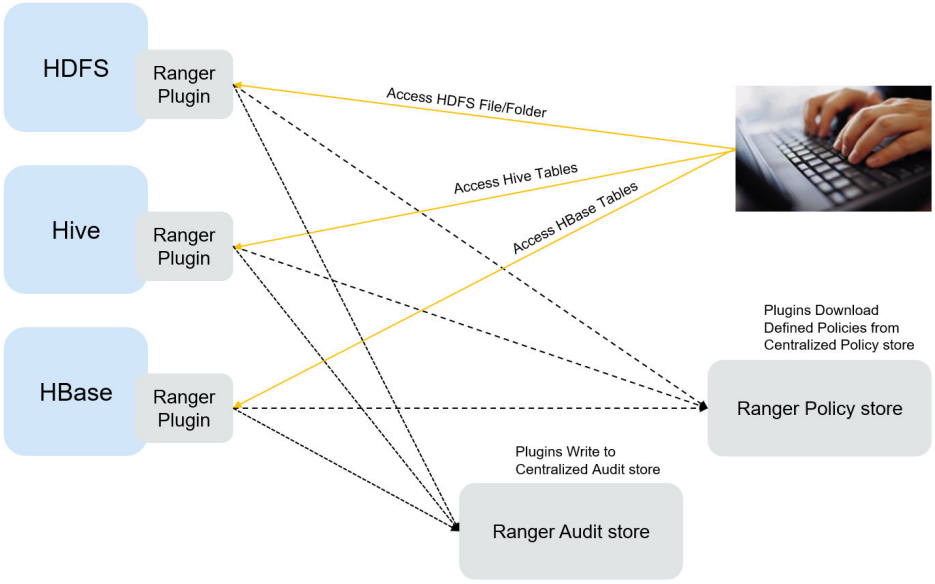
Ranger provides PABC-based authentication plug-ins for components to run on their servers. Ranger currently supports authentication for the following components like HDFS, YARN, Hive, HBase, Kafka, Storm, and Spark. More components will be supported in the future.

Ranger provides policy-based access control (PBAC) plug-ins to replace the original authentication plug-ins of the components. Ranger plug-ins are developed based on the authentication interface of the components. Users set permission policies for specified services on the Ranger web UI. Ranger plug-ins periodically update policies from the RangerAdmin and caches them in the local file of the component. When a client request needs to be authenticated, the Ranger plug-in matches the user carried in the request with the policy and then returns an accept or reject message.

Each time a component is started, the system checks whether the default Ranger service of the component exists. If the service does not exist, the system creates the Ranger service and adds a default policy for it. If a service is deleted by mistake, you can restart or restart the corresponding component service in rolling mode to restore the service. If the default policy is deleted by mistake, you can manually delete the service and then restart the component service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-93 Relationships between Ranger and other components



1.6.29 Spark

1.6.29.1 Spark Basic Principles

Description

Spark is an open-source parallel data processing framework. It helps you easily and quickly develop big data applications and perform offline processing, stream processing, and interactive analysis on data.

Spark provides a framework featuring fast computing, write, and interactive query. Spark has obvious advantages over Hadoop in terms of performance. Spark uses the in-memory computing mode to avoid I/O bottlenecks in scenarios where multiple tasks in a MapReduce workflow process the same dataset. Spark is implemented by using Scala programming language. Scala enables distributed datasets to be processed in a method that is the same as that of processing local data. In addition to interactive data analysis, Spark supports interactive data mining. Spark adopts in-memory computing, which facilitates iterative computing. By coincidence, iterative computing of the same data is a general problem facing data mining. In addition, Spark can run in Yarn clusters where Hadoop 2.0 is installed. The reason why Spark cannot only retain various features like MapReduce fault tolerance, data localization, and scalability but also ensure high performance and avoid busy disk I/Os is that a memory abstraction structure called Resilient Distributed Dataset (RDD) is created for Spark.

Original distributed memory abstraction, for example, key-value store and databases, supports small-granularity update of variable status. This requires backup of data or log updates to ensure fault tolerance. Consequently, a large amount of I/O consumption is brought about to data-intensive workflows. For the RDD, it has only one set of restricted APIs and only supports large-granularity update, for example, map and join. In this way, Spark only needs to record the transformation operation logs generated during data establishment to ensure

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

fault tolerance without recording a complete dataset. This data transformation link record is a source for tracing a data set. Generally, parallel applications apply the same computing process for a large dataset. Therefore, the limit to the mentioned large-granularity update is not large. As described in Spark theses, the RDD can function as multiple different computing frameworks, for example, programming models of MapReduce and Pregel. In addition, Spark allows you to explicitly make a data transformation process be persistent on hard disks. Data localization is implemented by allowing you to control data partitions based on the key value of each record. (An obvious advantage of this method is that two copies of data to be associated will be hashed in the same mode.) If memory usage exceeds the physical limit, Spark writes relatively large partitions into hard disks, thereby ensuring scalability.

Spark has the following features:

- Fast: Spark processes data 10 to 100 times faster than MapReduce in iterative computing, machine learning, and interactive query.
- Easy-to-use: Java, Scala, and Python can be used to simply and quickly compile parallel applications for processing massive amounts of data. Spark provides over 80 operators to help you compile parallel applications.
- Integration with Hadoop: Spark can directly run in a Hadoop cluster and read existing Hadoop data.

The Spark component of MRS has the following advantages:

- The Spark Streaming component of MRS supports real-time data processing rather than triggering as scheduled.
- The Spark component of MRS provides Structured Streaming and allows you to build streaming applications using the Dataset API. Spark supports exactly-once semantics and inner and outer joins for streams.
- The Spark component of MRS uses **pandas_udf** to replace the original user-defined functions (UDFs) in PySpark to process data, which reduces the processing duration by 60% to 90% (affected by specific operations).
- The Spark component of MRS also supports graph data processing and allows modeling using graphs during graph computing.
- Spark SQL of MRS is compatible with some Hive syntax (based on the 64 SQL statements of the Hive-Test-benchmark test set) and standard SQL syntax (based on the 99 SQL statements of the TPC-DS test set).

Architecture

Figure 1-94 describes the Spark architecture and **Table 1-24** lists the Spark modules.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-94 Spark architecture

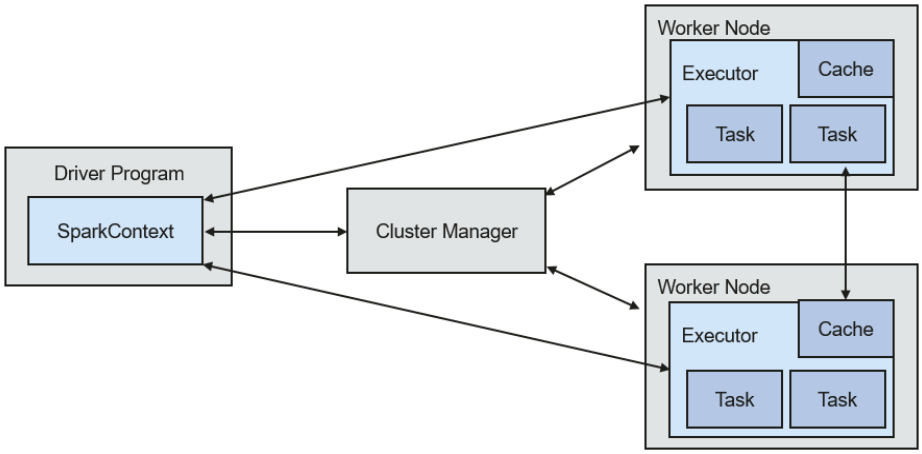


Table 1-24 Basic concepts

Module	Description
Cluster Manager	Cluster manager manages resources in the cluster. Spark supports multiple cluster managers, including Mesos, YARN, and the Standalone cluster manager that is delivered with Spark.
Application	Spark application. It consists of one Driver Program and multiple executors.
Deploy Mode	Deployment in cluster or client mode. In cluster mode, the driver runs on a node inside the cluster. In client mode, the driver runs on the client (outside the cluster).
Driver Program	The main process of the Spark application. It runs the **main()** function of an application and creates SparkContext. It is used for parsing applications, generating stages, and scheduling tasks to executors. Usually, SparkContext represents Driver Program.
Executor	A process started on a Work Node. It is used to execute tasks, and manage and process the data used in applications. A Spark application usually contains multiple executors. Each executor receives commands from the driver and executes one or multiple tasks.
Worker Node	A node that starts and manages executors and resources in a cluster.
Job	A job consists of multiple concurrent tasks. One action operator (for example, a collect operator) maps to one job.
Stage	Each job consists of multiple stages. Each stage is a task set, which is separated by Directed Acyclic Graph (DAG).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Description
Task	A task carries the computation unit of the service logics. It is the minimum working unit that can be executed on the Spark platform. An application can be divided into multiple tasks based on the execution plan and computation amount.

Spark Application Running Principle

Figure 1-95 shows the Spark application running architecture. The running process is as follows:

1. An application is running in the cluster as a collection of processes. Driver coordinates the running of the application.
2. To run an application, Driver connects to the cluster manager (such as Standalone, Mesos, and YARN) to apply for the executor resources, and start ExecutorBackend. The cluster manager schedules resources between different applications. Driver schedules DAGs, divides stages, and generates tasks for the application at the same time.
3. Then, Spark sends the code of the application (the code transferred to SparkContext, which is defined by JAR or Python) to an executor.
4. After all tasks are finished, the running of the user application is stopped.

Figure 1-95 Spark application running architecture

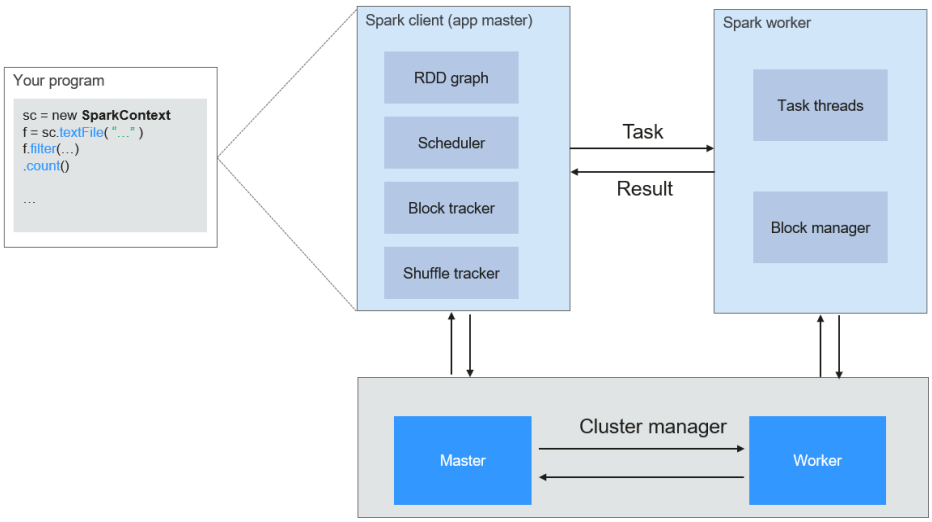
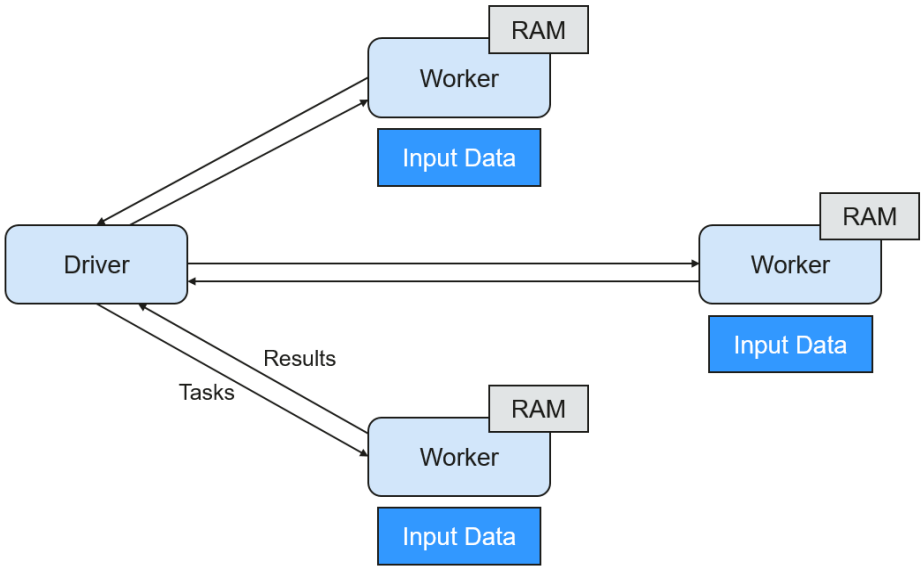


Figure 1-96 shows the Master and Worker modes adopted by Spark. A user submits an application on the Spark client, and then the scheduler divides a job into multiple tasks and sends the tasks to each Worker for execution. Each Worker reports the computation results to Driver (Master), and then the Driver aggregates and returns the results to the client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-96 Spark Master-Worker mode



Note the following about the architecture:

- Applications are isolated from each other.

  Each application has an independent executor process, and each executor starts multiple threads to execute tasks in parallel. Whether in terms of scheduling or task running on executors. Each driver independently schedules its own tasks. Different application tasks run on different JVMs, that is, different executors.
- Different Spark applications do not share data, unless data is stored in the external storage system such as HDFS.
- You are advised to deploy the Driver program in a location that is close to the Worker node because the Driver program schedules tasks in the cluster. For example, deploy the Driver program on the network where the Worker node is located.

Spark on YARN can be deployed in two modes:

- In YARN-cluster mode, the Spark driver runs inside an ApplicationMaster process which is managed by Yarn in the cluster. After the ApplicationMaster is started, the client can exit without interrupting service running.
- In Yarn-client mode, the driver is started in the client process, and the ApplicationMaster process is used only to apply for resources from the Yarn cluster.

Spark Streaming Principle

Spark Streaming is a real-time computing framework built on the Spark, which expands the capability for processing massive streaming data. Currently, Spark supports the following data processing methods:

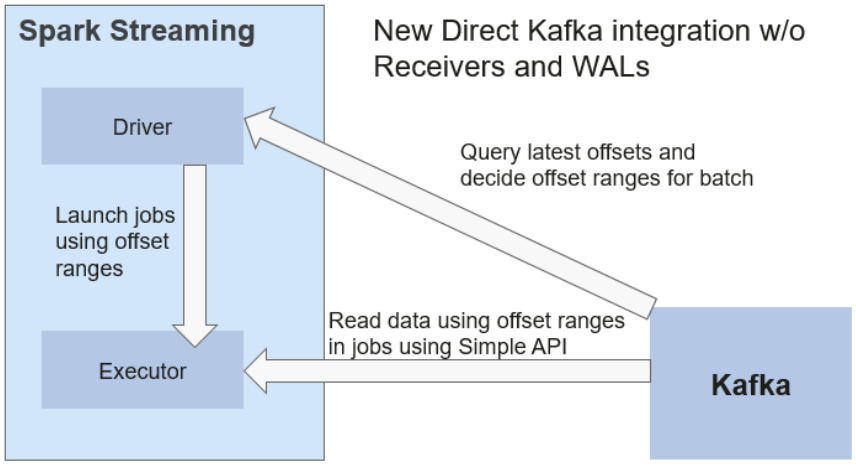
- Direct Streaming

  In Direct Streaming approach, Direct API is used to process data. Take Kafka Direct API as an example. Direct API provides offset location that each batch

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

range will read from, which is much simpler than starting a receiver to continuously receive data from Kafka and written data to write-ahead logs (WALs). Then, each batch job is running and the corresponding offset data is ready in Kafka. These offset information can be securely stored in the checkpoint file and read by applications that failed to start.

Figure 1-97 Data transmission through Direct Kafka API



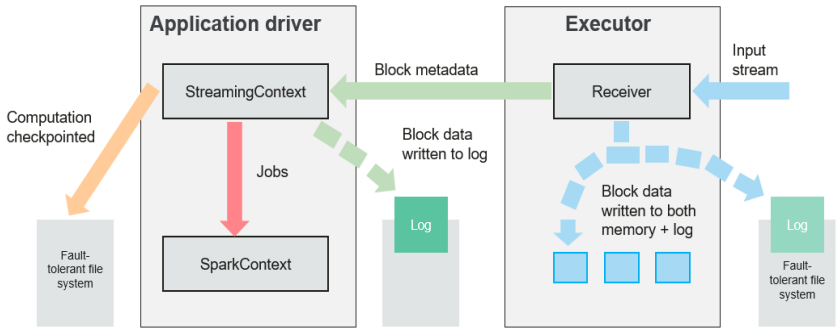
After the failure, Spark Streaming can read data from Kafka again and process the data segment. The processing result is the same no matter Spark Streaming fails or not, because the semantic is processed only once.

Direct API does not need to use the WAL and Receivers, and ensures that each Kafka record is received only once, which is more efficient. In this way, the Spark Streaming and Kafka can be well integrated, making streaming channels be featured with high fault-tolerance, high efficiency, and ease-of-use. Therefore, you are advised to use Direct Streaming to process data.

- Receiver

  When a Spark Streaming application starts (that is, when the driver starts), the related StreamingContext (the basis of all streaming functions) uses SparkContext to start the receiver to become a long-term running task. These receivers receive and save streaming data to the Spark memory for processing. **Figure 1-98** shows the data transfer lifecycle.

Figure 1-98 Data transfer lifecycle



a. Receive data (blue arrow).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Receiver divides a data stream into a series of blocks and stores them in the executor memory. In addition, after WAL is enabled, it writes data to the WAL of the fault-tolerant file system.

b. Notify the driver (green arrow).

   The metadata in the received block is sent to StreamingContext in the driver. The metadata includes:

   - Block reference ID used to locate the data position in the Executor memory.
   - Block data offset information in logs (if the WAL function is enabled).

c. Process data (red arrow).

   For each batch of data, StreamingContext uses block information to generate resilient distributed datasets (RDDs) and jobs. StreamingContext executes jobs by running tasks to process blocks in the executor memory.

d. Periodically set checkpoints (orange arrows).

   For fault tolerance, StreamingContext periodically sets checkpoints and saves them to external file systems.

Fault Tolerance

Spark and its RDDs allow seamless processing of any Worker node failure in the cluster. Spark Streaming is built on top of Spark. Therefore, the Worker node of Spark Streaming also has the same fault tolerance capability. However, Spark Streaming needs to run properly in case of long-time running. Therefore, Spark must be able to recover from faults through the driver process (main process that coordinates all Workers). This poses challenges to the Spark driver fault-tolerance because the Spark driver may be any user application implemented in any computation mode. However, Spark Streaming has internal computation architecture. That is, it periodically executes the same Spark computation in each batch data. Such architecture allows it to periodically store checkpoints to reliable storage space and recover them upon the restart of Driver.

For source data such as files, the Driver recovery mechanism can ensure zero data loss because all data is stored in a fault-tolerant file system such as HDFS. However, for other data sources such as Kafka and Flume, some received data is cached only in memory and may be lost before being processed. This is caused by the distribution operation mode of Spark applications. When the driver process fails, all executors running in the Cluster Manager, together with all data in the memory, are terminated. To avoid such data loss, the WAL function is added to Spark Streaming.

WAL is often used in databases and file systems to ensure persistence of any data operation. That is, first record an operation to a persistent log and perform this operation on data. If the operation fails, the system is recovered by reading the log and re-applying the preset operation. The following describes how to use WAL to ensure persistence of received data:

Receiver is used to receive data from data sources such as Kafka. As a long-time running task in Executor, Receiver receives data, and also confirms received data if supported by data sources. Received data is stored in the Executor memory, and Driver delivers a task to Executor for processing.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After WAL is enabled, all received data is stored to log files in the fault-tolerant file system. Therefore, the received data does not lose even if Spark Streaming fails. Besides, receiver checks correctness of received data only after the data is pre-written into logs. Data that is cached but not stored can be sent again by data sources after the driver restarts. These two mechanisms ensure zero data loss. That is, all data is recovered from logs or re-sent by data sources.

To enable the WAL function, perform the following operations:

- Set **streamingContext.checkpoint** to configure the checkpoint directory, which is an HDFS file path used to store streaming checkpoints and WALs.
- Set **spark.streaming.receiver.writeAheadLog.enable** of SparkConf to **true** (the default value is **false**).

After WAL is enabled, all receivers have the advantage of recovering from reliable received data. You are advised to disable the multi-replica mechanism because the fault-tolerant file system of WAL may also replicate the data.

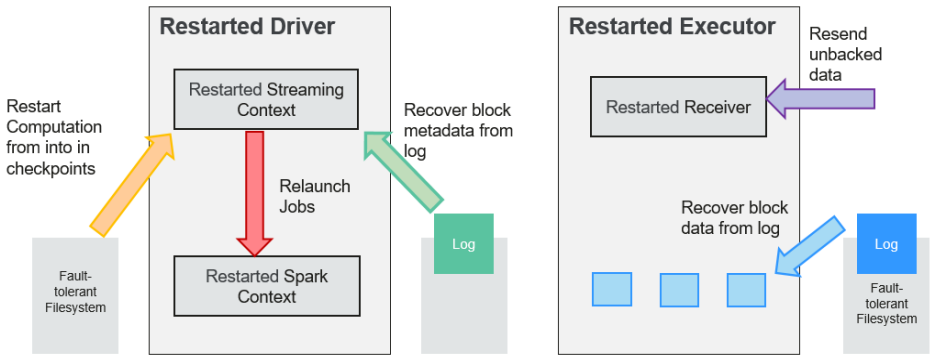
NOTE

The data receiving throughput is lowered after WAL is enabled. All data is written into the fault-tolerant file system. As a result, the write throughput of the file system and the network bandwidth for data replication may become the potential bottleneck. To solve this problem, create more receivers to increase the data receiving parallelism or use better hardware to improve the throughput of the fault-tolerant file system.

Recovery Process

When a failed driver is restarted, restart it as follows:

Figure 1-99 Computing recovery process



1. Recover computing. (Orange arrow)

   Use checkpoint information to restart Driver, reconstruct SparkContext and restart Receiver.

2. Recover metadata block. (Green arrow)

   This operation ensures that all necessary metadata blocks are recovered to continue the subsequent computing recovery.

3. Relaunch unfinished jobs. (Red arrow)

   Recovered metadata is used to generate RDDs and corresponding jobs for interrupted batch processing due to failures.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Read block data saved in logs. (Blue arrow)

   Block data is directly read from WALs during execution of the preceding jobs, and therefore all essential data reliably stored in logs is recovered.

5. Resend unconfirmed data. (Purple arrow)

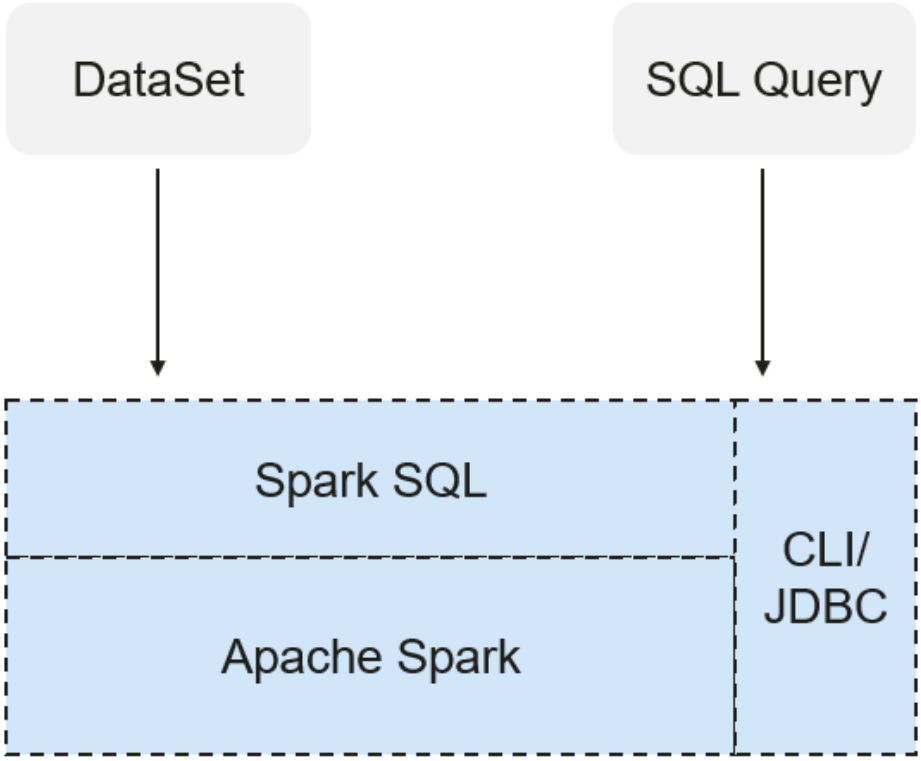
   Data that is cached but not stored to logs upon failures is re-sent by data sources, because the receiver does not confirm the data.

Therefore, by using WALs and reliable Receiver, Spark Streaming can avoid input data loss caused by Driver failures.

SparkSQL and DataSet Principle

SparkSQL

Figure 1-100 SparkSQL and DataSet



Spark SQL is a module for processing structured data. In Spark applications, SQL statements or DataSet APIs can be seamlessly used for querying structured data.

Spark SQL and DataSet also provide a universal method for accessing multiple data sources such as Hive, CSV, Parquet, ORC, JSON, and JDBC. These data sources also allow data interaction. Spark SQL reuses the Hive frontend processing logic and metadata processing module. With the Spark SQL, you can directly query existing Hive data.

In addition, Spark SQL also provides API, CLI, and JDBC APIs, allowing diverse accesses to the client.

Spark SQL Native DDL/DML

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In Spark 1.5, lots of Data Definition Language (DDL)/Data Manipulation Language (DML) commands are pushed down to and run on the Hive, causing coupling with the Hive and inflexibility such as unexpected error reports and results.

Spark 3.1.1 realizes command localization and replaces the Hive with Spark SQL Native DDL/DML to run DDL/DML commands. Additionally, the decoupling from the Hive is realized and commands can be customized.

DataSet

A DataSet is a strongly typed collection of domain-specific objects that can be transformed in parallel using functional or relational operations. Each Dataset also has an untyped view called a DataFrame, which is a Dataset of Row.

The DataFrame is a structured and distributed dataset consisting of multiple columns. The DataFrame is equal to a table in the relationship database or the DataFrame in the R/Python. The DataFrame is the most basic concept in the Spark SQL, which can be created by using multiple methods, such as the structured dataset, Hive table, external database or RDD.

Operations available on DataSets are divided into transformations and actions.

- A transformation operation can generate a new DataSet,
  for example, **map**, **filter**, **select**, and **aggregate (groupBy)**.
- An action operation can trigger computation and return results,
  for example, **count**, **show**, or write data to the file system.

You can use either of the following methods to create a DataSet:

- The most common way is by pointing Spark to some files on storage systems, using the **read** function available on a SparkSession.
  ```
  val people = spark.read.parquet("...").as[Person]  // Scala
  DataSet<Person> people = spark.read().parquet("...").as(Encoders.bean(Person.class));//Java
  ```
- You can also create a DataSet using the transformation operation available on an existing one.
  For example, apply the map operation on an existing DataSet to create a DataSet:
  ```
  val names = people.map(_.name) // In Scala: names is Dataset.
  Dataset<String> names = people.map((Person p) -> p.name, Encoders.STRING)); // Java
  ```

CLI and JDBCServer

In addition to programming APIs, Spark SQL also provides the CLI/JDBC APIs.

- Both **spark-shell** and **spark-sql** scripts can provide the CLI for debugging.
- JDBCServer provides JDBC APIs. External systems can directly send JDBC requests to calculate and parse structured data.

SparkSession Principle

SparkSession is a unified API for Spark programming and can be regarded as a unified entry for reading data. SparkSession provides a single entry point to perform many operations that were previously scattered across multiple classes, and also provides accessor methods to these older classes to maximize compatibility.

A SparkSession can be created using a builder pattern. The builder will automatically reuse the existing SparkSession if there is a SparkSession; or create

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

a SparkSession if it does not exist. During I/O transactions, the configuration item settings in the builder are automatically synchronized to Spark and Hadoop.

```
import org.apache.spark.sql.SparkSession
val sparkSession = SparkSession.builder
  .master("local")
  .appName("my-spark-app")
  .config("spark.some.config.option", "config-value")
  .getOrCreate()
```

- SparkSession can be used to execute SQL queries on data and return results as DataFrame.
  ```
  sparkSession.sql("select * from person").show
  ```
- SparkSession can be used to set configuration items during running. These configuration items can be replaced with variables in SQL statements.
  ```
  sparkSession.conf.set("spark.some.config", "abcd")
  sparkSession.conf.get("spark.some.config")
  sparkSession.sql("select ${spark.some.config}")
  ```
- SparkSession also includes a "catalog" method that contains methods to work with Metastore (data catalog). After this method is used, a dataset is returned, which can be run using the same Dataset API.
  ```
  val tables = sparkSession.catalog.listTables()
  val columns = sparkSession.catalog.listColumns("myTable")
  ```
- Underlying SparkContext can be accessed by SparkContext API of SparkSession.
  ```
  val sparkContext = sparkSession.sparkContext
  ```

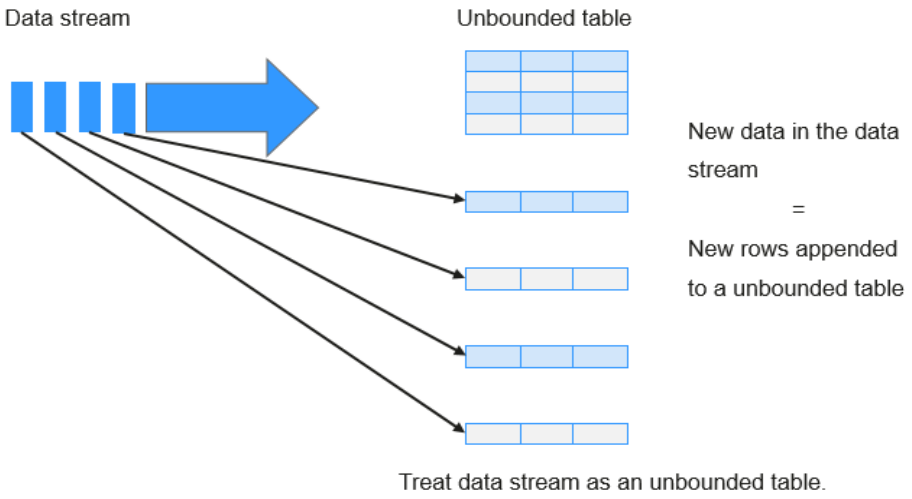
Structured Streaming Principle

Structured Streaming is a stream processing engine built on the Spark SQL engine. You can use the Dataset/DataFrame API in Scala, Java, Python, or R to express streaming aggregations, event-time windows, and stream-stream joins. If streaming data is incrementally and continuously produced, Spark SQL will continue to process the data and synchronize the result to the result set. In addition, the system ensures end-to-end exactly-once fault-tolerance guarantees through checkpoints and WALs.

The core of Structured Streaming is to take streaming data as an incremental database table. Similar to the data block processing model, the streaming data processing model applies query operations on a static database table to streaming computing, and Spark uses standard SQL statements for query, to obtain data from the incremental and unbounded table.

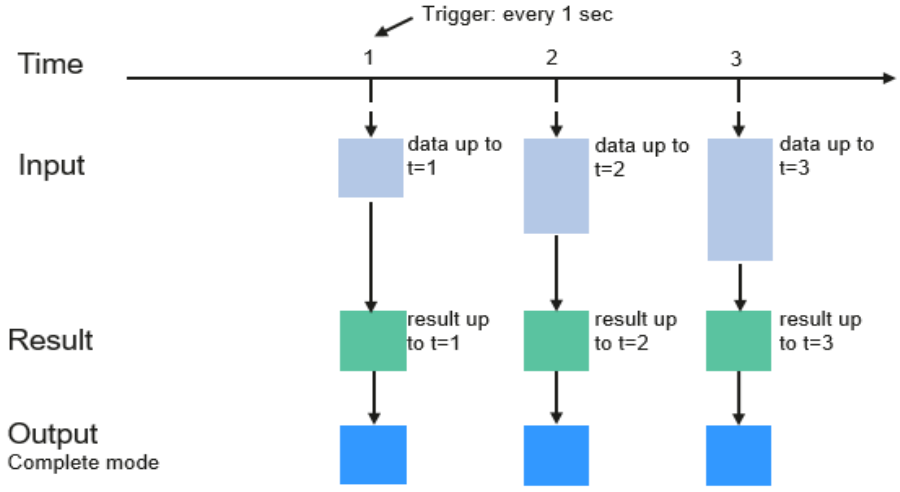
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-101 Unbounded table of Structured Streaming



Each query operation will generate a result table. At each trigger interval, updated data will be synchronized to the result table. Whenever the result table is updated, the updated result will be written into an external storage system.

Figure 1-102 Structured Streaming data processing model



Programming Model for Structured Streaming

Storage modes of Structured Streaming at the output phase are as follows:

- Complete Mode: The updated result sets are written into the external storage system. The write operation is performed by a connector of the external storage system.
- Append Mode: If an interval is triggered, only added data in the result table will be written into an external system. This is applicable only to the queries where existing rows in the result table are not expected to change.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Update Mode: If an interval is triggered, only updated data in the result table will be written into an external system, which is the difference between the Complete Mode and Update Mode.

Basic Concepts

- **RDD**

  Resilient Distributed Dataset (RDD) is a core concept of Spark. It indicates a read-only and partitioned distributed dataset. Partial or all data of this dataset can be cached in the memory and reused between computations.

  RDD Creation

  - An RDD can be created from the input of HDFS or other storage systems that are compatible with Hadoop.
  - A new RDD can be converted from a parent RDD.
  - An RDD can be converted from a collection of datasets through encoding.

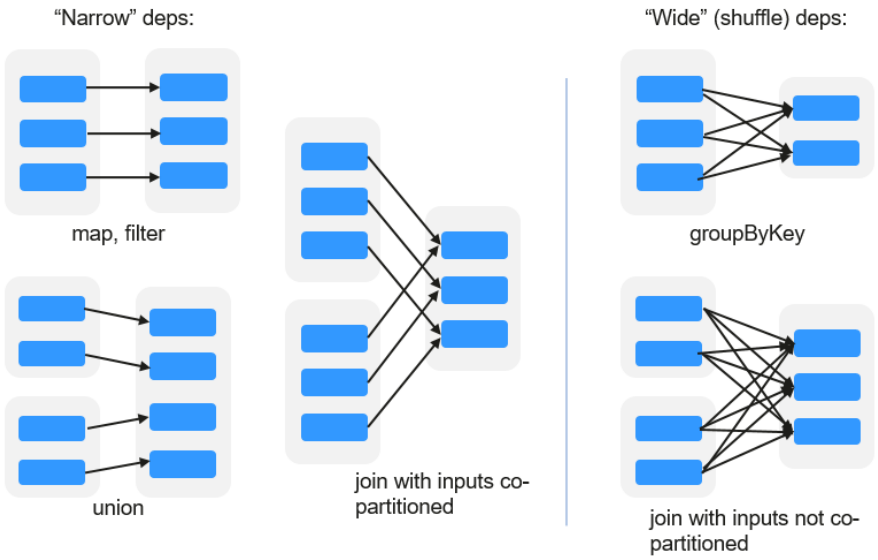
  RDD Storage

  - You can select different storage levels to store an RDD for reuse. (There are 11 storage levels to store an RDD.)
  - By default, the RDD is stored in the memory. When the memory is insufficient, the RDD overflows to the disk.

- **RDD Dependency**

  The RDD dependency includes the narrow dependency and wide dependency.

  Figure 1-103 RDD dependency

  - **Narrow dependency**: Each partition of the parent RDD is used by at most one partition of the child RDD.
  - **Wide dependency**: Partitions of the child RDD depend on all partitions of the parent RDD.

  The narrow dependency facilitates the optimization. Logically, each RDD operator is a fork/join (the join is not the join operator mentioned above but the barrier used to synchronize multiple concurrent tasks); fork the RDD to

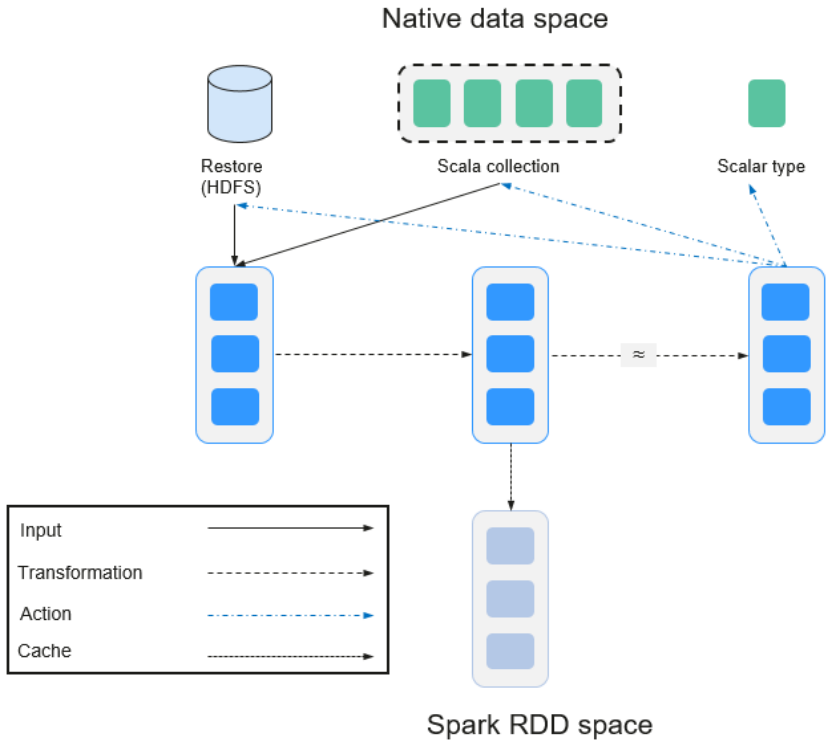
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

each partition, and then perform the computation. After the computation, join the results, and then perform the fork/join operation on the next RDD operator. It is uneconomical to directly translate the RDD into physical implementation. The first is that every RDD (even intermediate result) needs to be physicalized into memory or storage, which is time-consuming and occupies much space. The second is that as a global barrier, the join operation is very expensive and the entire join process will be slowed down by the slowest node. If the partitions of the child RDD narrowly depend on that of the parent RDD, the two fork/join processes can be combined to implement classic fusion optimization. If the relationship in the continuous operator sequence is narrow dependency, multiple fork/join processes can be combined to reduce a large number of global barriers and eliminate the physicalization of many RDD intermediate results, which greatly improves the performance. This is called pipeline optimization in Spark.

- **Transformation and Action (RDD Operations)**

  Operations on RDD include transformation (the return value is an RDD) and action (the return value is not an RDD). **Figure 1-104** shows the RDD operation process. The transformation is lazy, which indicates that the transformation from one RDD to another RDD is not immediately executed. Spark only records the transformation but does not execute it immediately. The real computation is started only when the action is started. The action returns results or writes the RDD data into the storage system. The action is the driving force for Spark to start the computation.

  Figure 1-104 RDD operation



The data and operation model of RDD are quite different from those of Scala.

```
val file = sc.textFile("hdfs://...")
val errors = file.filter(_.contains("ERROR"))
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
errors.cache()
errors.count()
```

a. The textFile operator reads log files from the HDFS and returns files (as an RDD).
b. The filter operator filters rows with **ERROR** and assigns them to errors (a new RDD). The filter operator is a transformation.
c. The cache operator caches errors for future use.
d. The count operator returns the number of rows of errors. The count operator is an action.

Transformation includes the following types:

- The RDD elements are regarded as simple elements.

  The input and output has the one-to-one relationship, and the partition structure of the result RDD remains unchanged, for example, map.

  The input and output has the one-to-many relationship, and the partition structure of the result RDD remains unchanged, for example, flatMap (one element becomes a sequence containing multiple elements after map and then flattens to multiple elements).

  The input and output has the one-to-one relationship, but the partition structure of the result RDD changes, for example, union (two RDDs integrates to one RDD, and the number of partitions becomes the sum of the number of partitions of two RDDs) and coalesce (partitions are reduced).

  Operators of some elements are selected from the input, such as filter, distinct (duplicate elements are deleted), subtract (elements only exist in this RDD are retained), and sample (samples are taken).
- The RDD elements are regarded as key-value pairs.

  Perform the one-to-one calculation on the single RDD, such as mapValues (the partition mode of the source RDD is retained, which is different from map).

  Sort the single RDD, such as sort and partitionBy (partitioning with consistency, which is important to the local optimization).

  Restructure and reduce the single RDD based on key, such as groupByKey and reduceByKey.

  Join and restructure two RDDs based on the key, such as join and cogroup.

  NOTE

  The later three operations involving sorting are called shuffle operations.

Action includes the following types:

- Generate scalar configuration items, such as **count** (the number of elements in the returned RDD), **reduce**, **fold/aggregate** (the number of scalar configuration items that are returned), and **take** (the number of elements before the return).
- Generate the Scala collection, such as **collect** (import all elements in the RDD to the Scala collection) and **lookup** (look up all values corresponding to the key).
- Write data to the storage, such as **saveAsTextFile** (which corresponds to the preceding **textFile**).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

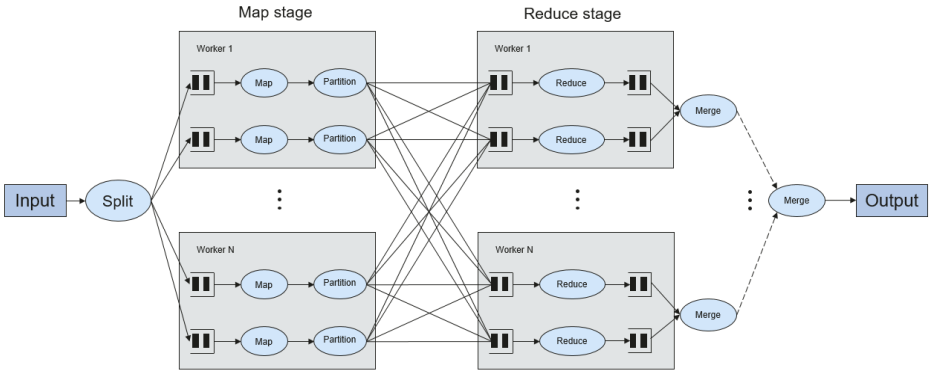
  - Check points, such as the **checkpoint** operator. When Lineage is quite long (which occurs frequently in graphics computation), it takes a long period of time to execute the whole sequence again when a fault occurs. In this case, checkpoint is used as the check point to write the current data to stable storage.

- **Shuffle**

  Shuffle is a specific phase in the MapReduce framework, which is located between the Map phase and the Reduce phase. If the output results of Map are to be used by Reduce, the output results must be hashed based on a key and distributed to each Reducer. This process is called Shuffle. Shuffle involves the read and write of the disk and the transmission of the network, so that the performance of Shuffle directly affects the operation efficiency of the entire program.

  The figure below shows the entire process of the MapReduce algorithm.

  Figure 1-105 Algorithm process



  Shuffle is a bridge connecting data. The following describes the implementation of shuffle in Spark.

  Shuffle divides a job of Spark into multiple stages. The former stages contain one or more ShuffleMapTasks, and the last stage contains one or more ResultTasks.

- **Spark Application Structure**

  The Spark application structure includes the initialized SparkContext and the main program.

  - Initialized SparkContext: constructs the operating environment of the Spark Application.

    Constructs the SparkContext object. The following is an example:

    ```
    new SparkContext(master, appName, [SparkHome], [jars])
    ```

    Parameter description:

    master: indicates the link string. The link modes include local, YARN-cluster, and YARN-client.

    appName: indicates the application name.

    SparkHome: indicates the directory where Spark is installed in the cluster.

    jars: indicates the code and dependency package of an application.

  - Main program: processes data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Spark Shell Commands**

  The basic Spark shell commands support the submission of Spark applications. The Spark shell commands are as follows:

```
./bin/spark-submit \
  --class <main-class> \
  --master <master-url> \
  ... # other options
  <application-jar> \
  [application-arguments]
```

  Parameter description:

  --class: indicates the name of the class of a Spark application.

  --master: indicates the master to which the Spark application links, such as YARN-client and YARN-cluster.

  application-jar: indicates the path of the JAR file of the Spark application.

  application-arguments: indicates the parameter required to submit the Spark application. This parameter can be left blank.

- **Spark JobHistory Server**

  The Spark web UI is used to monitor the details in each phase of the Spark framework of a running or historical Spark job and provide the log display, which helps users to develop, configure, and optimize the job in more fine-grained units.

1.6.29.2 Spark HA Solution

Spark Multi-Active Instance HA Principles and Implementation Solution

Based on existing JDBCServer in the community, multi-active-instance mode is used to achieve HA. In this mode, multiple JDBCServers coexist in the cluster and the client can randomly connect any JDBCServer to perform service operations. When one or multiple JDBCServers stop working, a client can connect to another normal JDBCServer.

Compared with active/standby HA mode, multi-active instance mode has following advantages:

- In active/standby HA, when the active/standby switchover occurs, the unavailable period cannot be controlled by JDBCServer, but it depends on Yarn service resources.
- In Spark, the Thrift JDBC similar to HiveServer2 provides services and users access services through Beeline and JDBC API. Therefore, the processing capability of the JDBCServer cluster depends on the single-point capability of the primary server, and the scalability is insufficient.

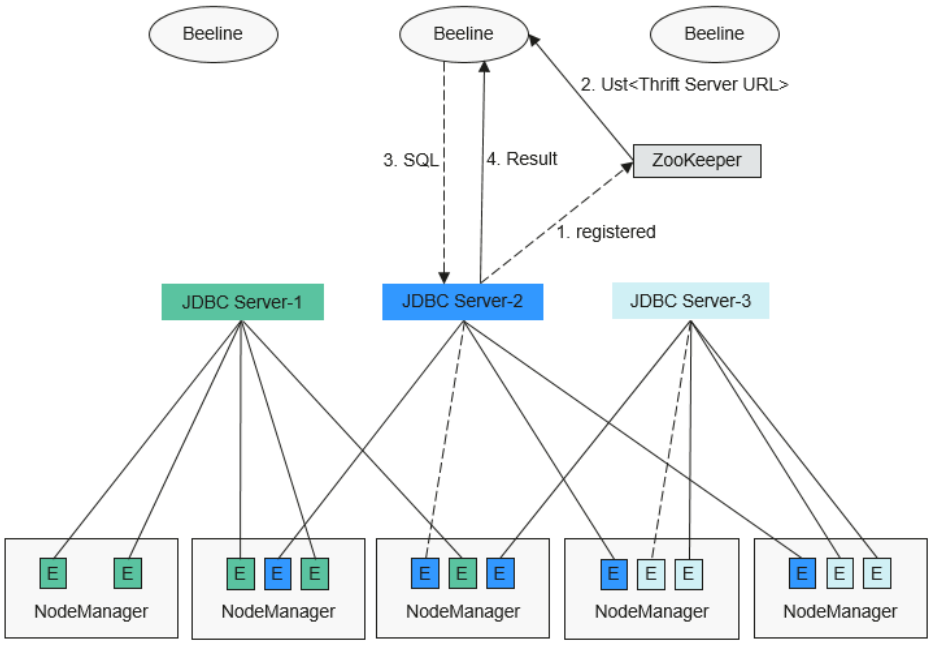
The multi-active instance HA mode not only can prevent service interruption caused by switchover, but also enables cluster scale-out to improve high concurrency.

- **Implementation**

  The following figure shows the basic principle of multi-active instance HA of Spark JDBCServer.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-106 Spark JDBCServer HA



1. When a JDBCServer is started, it registers with ZooKeeper by writing node information in a specified directory. Node information includes the instance IP address, port number, version, and serial number.
2. To connect to JDBCServer, the client must specify the namespace, which is the directory of JDBCServer instances in ZooKeeper. During the connection, a JDBCServer instance is randomly selected from the specified namespace.
3. After the connection succeeds, the client sends SQL statements to JDBCServer.
4. JDBCServer executes received SQL statements and returns results to the client.

If multi-active instance HA of Spark JDBCServer is enabled, all JDBCServer instances are independent and equivalent. When one JDBCServer instance is interrupted during upgrade, other JDBCServer instances can accept the connection request from the client.

The rules below must be followed in the multi-active instance HA of Spark JDBCServer.

- If a JDBCServer instance exits abnormally, no other instance will take over the sessions and services running on the abnormal instance.
- When the JDBCServer process is stopped, corresponding nodes are deleted from ZooKeeper.
- The client randomly selects the server, which may result in uneven session allocation caused by random distribution of policy results, and finally result in load imbalance of instances.
- After the instance enters the maintenance mode (in which no new connection requests from clients are accepted), services running on the instance may fail when the decommissioning times out.

- **URL Connection**
  - Multi-active instance mode

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In multi-active instance mode, the client reads content from the ZooKeeper node and connects to JDBCServer. The connection strings are list below.

- Security mode:

  If Kinit authentication is enabled, the JDBCURL is as follows:
  jdbc:hive2://<*zkNode1_IP*>:<*zkNode1_Port*>,<*zkNode2_IP*>:<*zkNode2_Port*>,<*zkNode3_IP*>:<*zkNode3_Port*>**/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.**<*System domain name*>**@**<*System domain name*>;

  NOTE

  - In the above JDBCURL, **<zkNode_IP>:<zkNode_Port>** indicates the ZooKeeper URL. Use commas (,) to separate multiple URLs.

    Example: 192.168.81.37:2181,192.168.195.232:2181,192.168.169.84:2181.
  - **sparkthriftserver2x** indicates the ZooKeeper directory, where a random JDBCServer instance is connected to the client.

  For example, when you use Beeline client to connect JDBCServer, run the following command:

  sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://**<*zkNode1_IP*>:<*zkNode1_Port*>,<*zkNode2_IP*>:<*zkNode2_Port*>,<*zkNode3_IP*>:<*zkNode3_Port*>**/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.**<*System domain name*>**@**<*System domain name*>**;"**

  If Keytab authentication is enabled, the JDBCURL is as follows:
  jdbc:hive2://<*zkNode1_IP*>:<*zkNode1_Port*>,<*zkNode2_IP*>:<*zkNode2_Port*>,<*zkNode3_IP*>:<*zkNode3_Port*>/**;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.**<*System domain name*>@<*System domain name*>**;user.principal=**<*principal_name*>**;user.keytab=**<*path_to_keytab*>

  In the above URL, <*principal_name*> indicates the principal of the Kerberos user, for example, **test@**<*System domain name*>; <*path_to_keytab*> indicates the Keytab file path corresponding to <*principal_name*>, for example, **/opt/auth/test/user.keytab**.

- Common mode:
  jdbc:hive2://<*zkNode1_IP*>:<*zkNode1_Port*>,<*zkNode2_IP*>:<*zkNode2_Port*>,<*zkNode3_IP*>:<*zkNode3_Port*>**/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;**

  For example, when you use Beeline client, in normal mode, for connection, run the following command:

  sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://**<*zkNode1_IP*>:<*zkNode1_Port*>,<*zkNode2_IP*>:<*zkNode2_Port*>,<*zkNode3_IP*>:<*zkNode3_Port*>**/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;"**

– Non-multi-active instance mode

  In this mode, a client connects to a specified JDBCServer node. Compared with multi-active instance mode, the connection string in this mode does not contain **serviceDiscoveryMode** and **zooKeeperNamespace** parameters about ZooKeeper.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, when you use Beeline client, in security mode, to connect JDBCServer in non-multi-active instance mode, run the following command:

sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://** *<server_IP>:<server_Port>*/**;user.principal=spark/hadoop.** *<System domain name>@<System domain name>*;**saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.** *<System domain name>@<System domain name>*;"

NOTE

- In the above command, **<server_IP>:<server_Port>** indicates the URL of the specified JDBCServer node.
- **CLIENT_HOME** indicates the client path.

Except the connection method, other operations of JDBCServer API in the two modes are the same.

Spark Multi-Tenant HA

In the JDBCServer multi-active instance solution, JDBCServer uses the Yarn-client mode, but there is only one Yarn resource queue available. To solve this resource limitation problem, the multi-tenant mode is introduced.

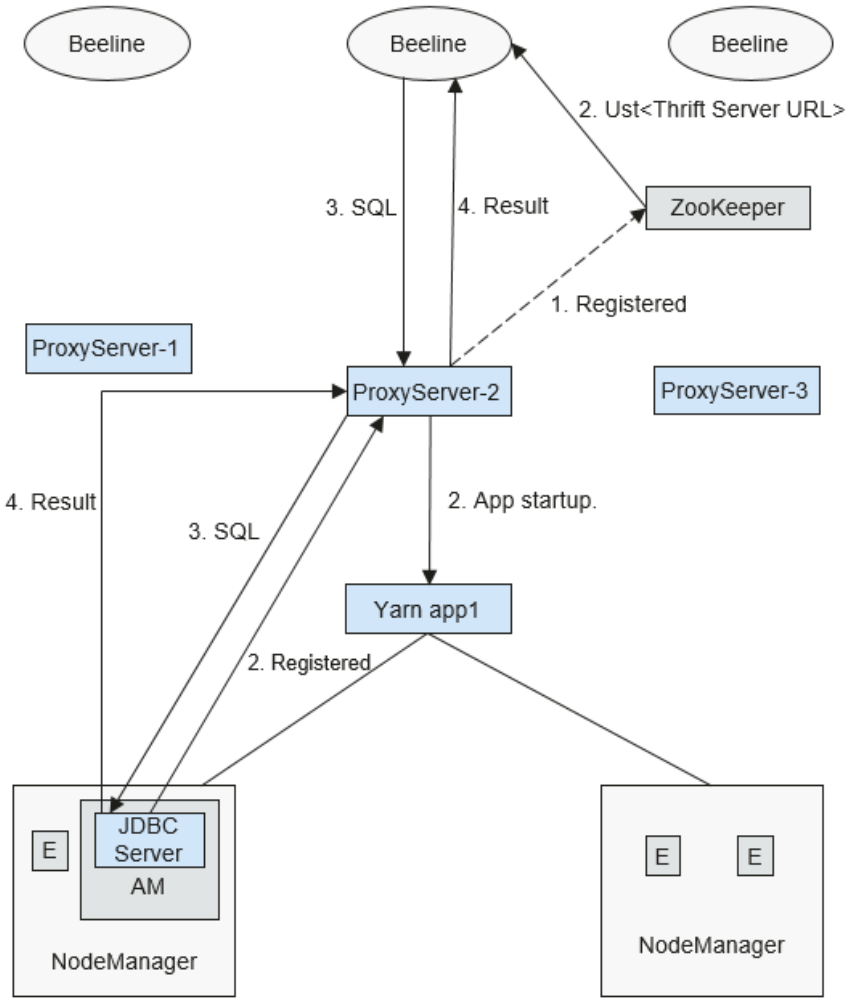
In multi-tenant mode, JDBCServers are bound with tenants. Each tenant corresponds to one or more JDBCServers, and a JDBCServer provides services for only one tenant. Different tenants can be configured with different Yarn queues to implement resource isolation. In addition, JDBCServer can be dynamically started as required to avoid resource waste.

- **Implementation**

  Figure 1-107 shows the HA solution of the multi-tenant mode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-107 Multi-tenant mode of Spark JDBCServer



a. When ProxyServer is started, it registers with ZooKeeper by writing node information in a specified directory. Node information includes the instance IP address, port number, version, and serial number.

   NOTE

   In multi-tenant mode, the JDBCServer instance refers to the ProxyServer (JDBCServer proxy).

b. To connect to ProxyServer, the client must specify a namespace, which is the directory of the ProxyServer instance where you want to access ZooKeeper. When the client connects to the ProxyServer, a random instance under the namespace is selected for connection. For details about the URL, see **URL Connection Overview**.

c. After the client successfully connects to the ProxyServer, which first checks whether the JDBCServer of a tenant exists. If yes, Beeline connects the JDBCServer. If no, a new JDBCServer is started in Yarn-cluster mode. After the startup of JDBCServer, ProxyServer obtains the IP address of the JDBCServer and establishes the connection between Beeline and JDBCServer.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

d. The client sends SQL statements to ProxyServer, which forwards statements to the connected JDBCServer. JDBCServer returns the results to ProxyServer, which then returns the results to the client.

In the multi-active instance HA mode, all instances are independent and equivalent. If one instance is interrupted during upgrade, other instances can accept the connection request from the client.

- **URL Connection Overview**
  - Multi-tenant mode

    In multi-tenant mode, the client reads content from the ZooKeeper node and connects to ProxyServer. The connection strings are list below.

    - Security mode:

      If Kinit authentication is enabled, the client URL is as follows:

      jdbc:hive2://
*<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.***<System domain name>**@**<System domain name>**;

      NOTE

      - In the above URL, **<zkNode_IP>:<zkNode_Port>** indicates the ZooKeeper URL. Use commas (,) to separate multiple URLs.

        Example:
        192.168.81.37:2181,192.168.195.232:2181,192.168.169.84:2181.
      - **sparkthriftserver2x** indicates the ZooKeeper directory, where a random JDBCServer instance is connected to the client.

      For example, when you use Beeline client for connection, run the following command:

      sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://***<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.***<System domain name>***@**<System domain name>***;"**

      If Keytab authentication is enabled, the URL is as follows:

      jdbc:hive2://
*<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.***<System domain name>**@**<System domain name>***;user.principal=***<principal_name>***;user.keytab=***<path_to_keytab>**

      In the above URL, *<principal_name>* indicates the principal of the Kerberos user, for example, **test@***<System domain name>*; *<path_to_keytab>* indicates the Keytab file path corresponding to *<principal_name>*, for example, **/opt/auth/test/user.keytab**.

    - Common mode:

      jdbc:hive2://
<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=**sparkthriftserver2x**;

      For example, run the following command when you use Beeline client for connection in normal mode:

      sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://***<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkN*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ode3_IP>:<zkNode3_Port>/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;"

- Non-multi-tenant mode

  In non-multi-tenant mode, a client connects to a specified JDBCServer node. Compared with multi-tenant instance mode, the connection string in this mode does not contain **serviceDiscoveryMode** and **zooKeeperNamespace** parameters about ZooKeeper.

  For example, when you use Beeline client to connect JDBCServer in non-multi-tenant instance mode, run the following command:

  sh *CLIENT_HOME*/spark/bin/beeline -u "jdbc:hive2://*<server_IP>:<server_Port>*/;user.principal=spark/hadoop.*<System domain name>*@*<System domain name>*;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.*<System domain name>*@*<System domain name>*;"

  NOTE

  - In the above command, **<server_IP>:<server_Port>** indicates the URL of the specified JDBCServer node.
  - **CLIENT_HOME** indicates the client path.

  Except the connection method, other operations of JDBCServer API in multi-tenant mode and non-multi-tenant mode are the same.

  Specifying a Tenant

  Generally, the client submitted by a user connects to the default JDBCServer of the tenant to which the user belongs. If you want to connect the client to the JDBCServer of a specified tenant, add the **--hiveconf mapreduce.job.queuename** parameter.

  If you use Beeline client for connection, run the following command (**aaa** is the tenant name):

  beeline --hiveconf mapreduce.job.queuename=aaa -u 'jdbc:hive2://192.168.39.30:2181,192.168.40.210:2181,192.168.215.97:2181;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.*<System domain name>*@*<System domain name>*;'

1.6.29.3 Relationship Between Spark and Other Components

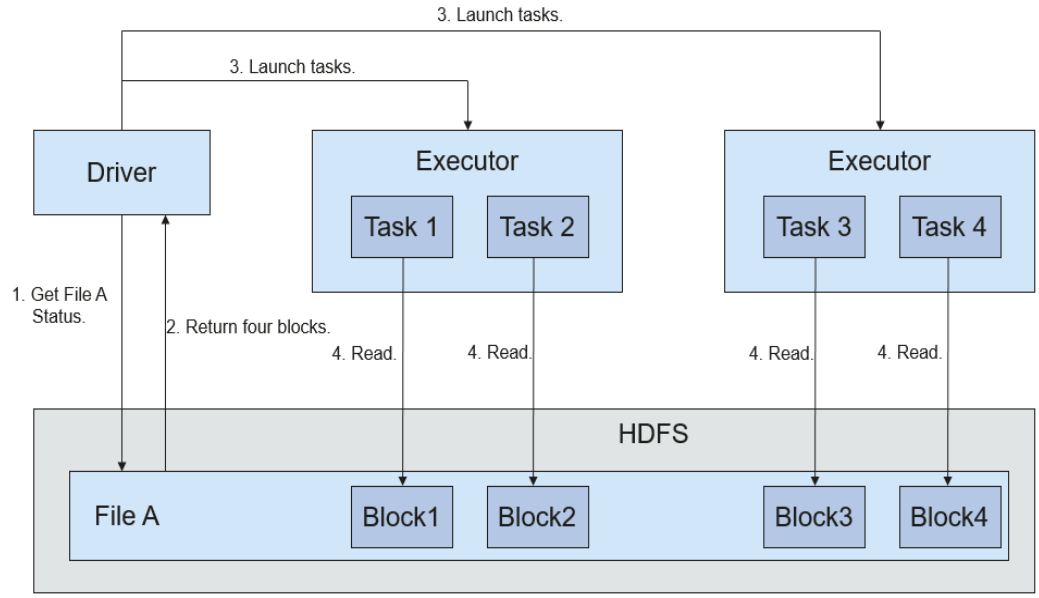
Relationship Between Spark and HDFS

Data computed by Spark comes from multiple data sources, such as local files and HDFS. Most data computed by Spark comes from the HDFS. The HDFS can read data in large scale for parallel computing. After being computed, data can be stored in the HDFS.

Spark involves Driver and Executor. Driver schedules tasks and Executor runs tasks.

Figure 1-108 shows the process of reading a file.

Figure 1-108 File reading process

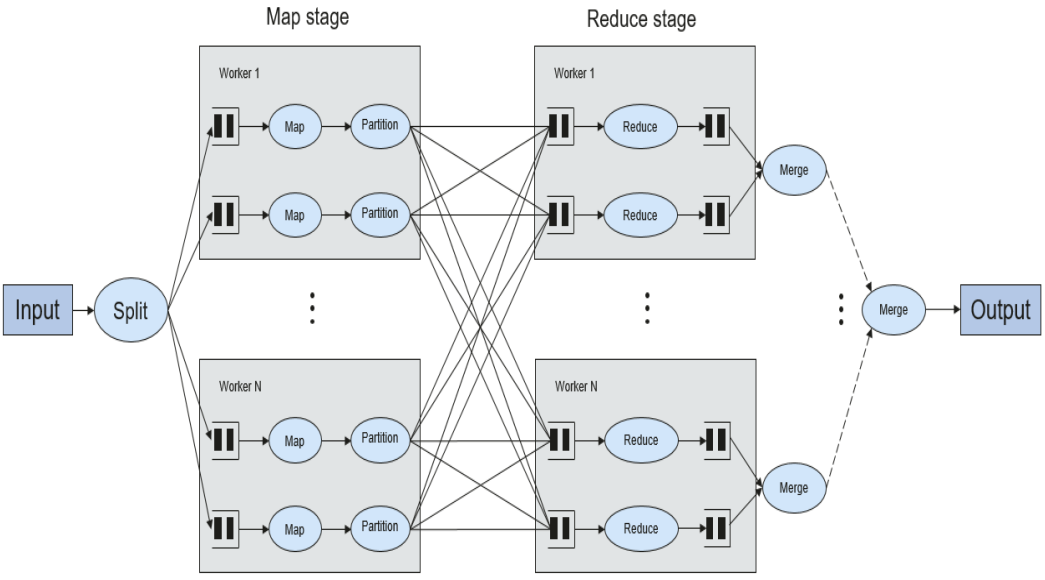


The file reading process is as follows:

1. Driver interconnects with the HDFS to obtain the information of File A.
2. The HDFS returns the detailed block information about this file.
3. Driver sets a parallel degree based on the block data amount, and creates multiple tasks to read the blocks of this file.
4. Executor runs the tasks and reads the detailed blocks as part of the Resilient Distributed Dataset (RDD).

Figure 1-109 shows the process of writing data to a file.

Figure 1-109 File writing process



The file writing process is as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Driver creates a directory where the file is to be written.
2. Based on the RDD distribution status, the number of tasks related to data writing is computed, and these tasks are sent to Executor.
3. Executor runs these tasks, and writes the RDD data to the directory created in 1.

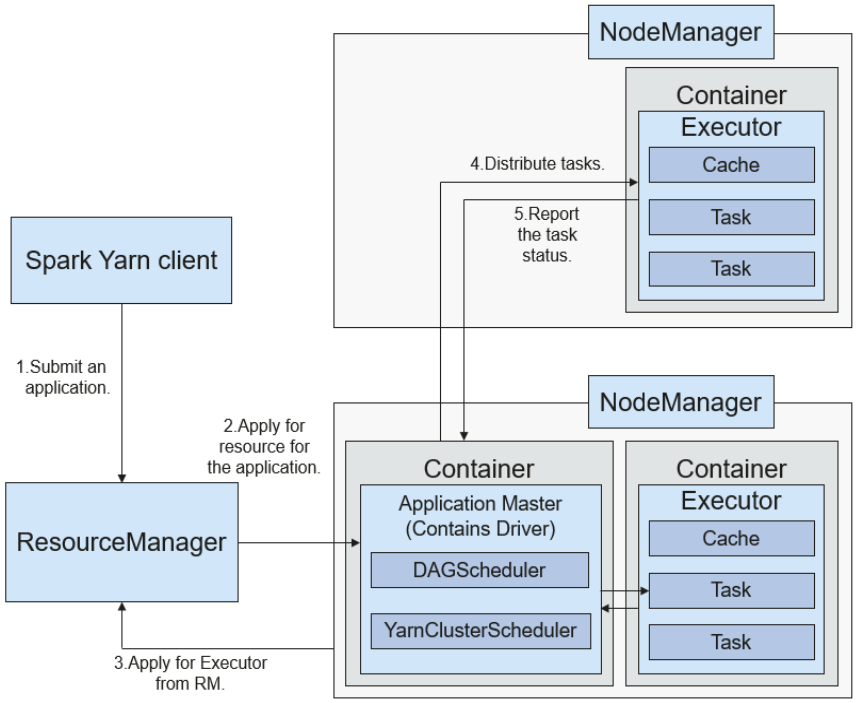
Relationship Between Spark and Yarn

The Spark computing and scheduling can be implemented using Yarn mode. Spark enjoys the computing resources provided by Yarn clusters and runs tasks in a distributed way. Spark on Yarn has two modes: Yarn-cluster and Yarn-client.

- Yarn-cluster mode

  Figure 1-110 shows the running framework of Spark on Yarn-cluster.

  Figure 1-110 Spark on Yarn-cluster operation framework



  Spark on Yarn-cluster implementation process:

  a. The client generates the application information, and then sends the information to ResourceManager.
  b. ResourceManager allocates the first container (ApplicationMaster) to SparkApplication and starts driver on the container.
  c. ApplicationMaster applies for resources from ResourceManager to run the container.

     ResourceManager allocates the container to ApplicationMaster, which communicates with NodeManager, and starts the executor in the obtained container. After the executor is started, it registers with the driver and applies for tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

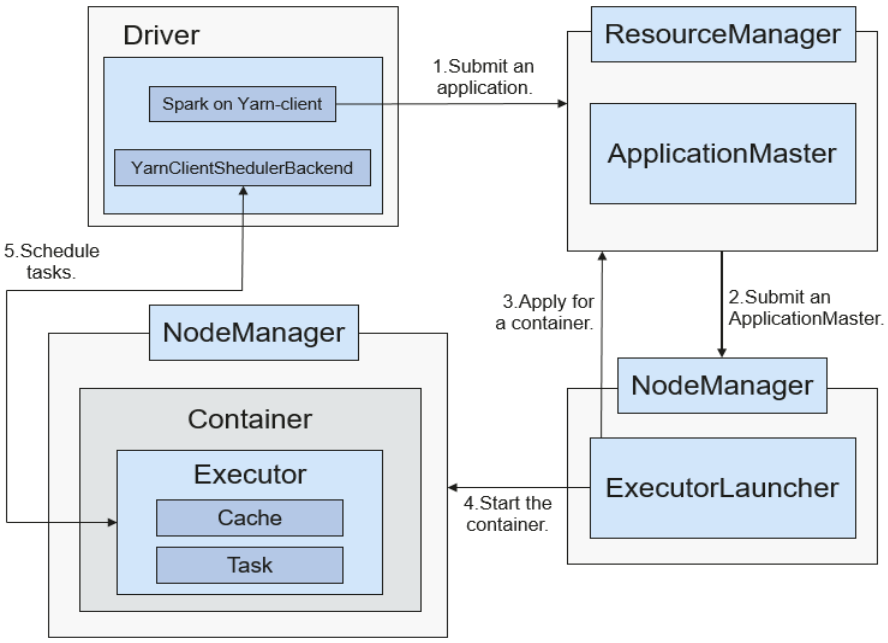
d. The driver allocates tasks to the executor.

e. The executor runs tasks and reports the operating status to the driver.

- Yarn-client mode

  Figure 1-111 shows the running framework of Spark on Yarn-cluster.

  Figure 1-111 Spark on Yarn-client operation framework



Spark on Yarn-client implementation process:

NOTE

In Yarn-client mode, Driver is deployed on the client and started on the client. In Yarn-client mode, the client of the earlier version is incompatible. You are advised to use the Yarn-cluster mode.

a. The client sends the Spark application request to ResourceManager, then ResourceManager returns the results. The results include information such as Application ID and the maximum and minimum available resources. The client packages all information required to start ApplicationMaster, and sends the information to ResourceManager.

b. After receiving the request, ResourceManager finds a proper node for ApplicationMaster and starts it on this node. ApplicationMaster is a role in Yarn, and the process name in Spark is ExecutorLauncher.

c. ApplicationMaster can apply for a series of containers to run tasks from ResourceManager.

d. After receiving the newly allocated container list (from ResourceManager), ApplicationMaster sends information to the related NodeManagers to start the containers.

   ResourceManager allocates the containers to ApplicationMaster, which communicates with the related NodeManagers, and starts the executors in the obtained containers. After the executors are started, it registers with drivers and applies for tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

Running containers are not suspended and resources are not released.

e. The drivers allocate tasks to the executors. The executor executes tasks and reports the operating status to the driver.

1.6.29.4 Spark Enhanced Open Source Features

Optimizing SQL Query of Data of Multiple Sources

Enterprises usually store massive data, such as from various databases and warehouses, for management and information collection. However, diversified data sources, hybrid dataset structures, and scattered data storage lower query efficiency.

The open source Spark only supports simple filter pushdown during querying of multi-source data. The SQL engine performance is deteriorated due of a large amount of unnecessary data transmission. The pushdown function is enhanced, so that **aggregate**, complex **projection**, and complex **predicate** can be pushed to data sources, reducing unnecessary data transmission and improving query performance.

Only the JDBC data source supports pushdown of query operations, such as **aggregate**, **projection**, **predicate**, **aggregate over inner join**, and **aggregate over union all**. All pushdown operations can be enabled based on your requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-25 Enhanced query of cross-source query

Module	Before Enhancement	After Enhancement
aggregate	The pushdown of **aggregate** is not supported.	• Aggregation functions including **sum**, **avg**, **max**, **min**, and **count** are supported. Example: select count(*) from table • Internal expressions of aggregation functions are supported. Example: select sum(a+b) from table • Calculation of aggregation functions is supported. Example: select avg(a) + max(b) from table • Pushdown of **having** is supported. Example: select sum(a) from table where a>0 group by b having sum(a)>10 • Pushdown of some functions is supported. Pushdown of lines in mathematics, time, and string functions, such as **abs()**, **month()**, and **length()** are supported. In addition to the preceding built-in functions, you can run the **SET** command to add functions supported by data sources. Example: select sum(abs(a)) from table • Pushdown of **limit** and **order by** after **aggregate** is supported. However, the pushdown is not supported in Oracle, because Oracle does not support **limit**. Example: select sum(a) from table where a>0 group by b order by sum(a) limit 5
projection	Only pushdown of simple **projection** is supported. Example: select a, b from table	• Complex expressions can be pushed down. Example: select (a+b)*c from table • Some functions can be pushed down. For details, see the description below the table. Example: select length(a)+abs(b) from table • Pushdown of **limit** and **order by** after **projection** is supported. Example: select a, b+c from table order by a limit 3

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Before Enhancement	After Enhancement
predicate	Only simple filtering with the column name on the left of the operator and values on the right is supported. Example: select * from table where a>0 or b in ("aaa", "bbb")	• Complex expression pushdown is supported. Example: select * from table where a +b>c*d or a/c in (1, 2, 3) • Some functions can be pushed down. For details, see the description below the table. Example: select * from table where length(a)>5
aggregate over inner join	Related data from the two tables must be loaded to Spark. The join operation must be performed before the **aggregate** operation.	The following functions are supported: • Aggregation functions including **sum**, **avg**, **max**, **min**, and **count** are supported. • All **aggregate** operations can be performed in a same table. The **group by** operations can be performed on one or two tables and only inner join is supported. The following scenarios are not supported: • **aggregate** cannot be pushed down from both the left- and right-join tables. • **aggregate** contains operations, for example, sum(a+b). • **aggregate** operations, for example, sum(a)+min(b).
aggregate over union all	Related data from the two tables must be loaded to Spark. **union** must be performed before **aggregate**.	Supported scenarios: Aggregation functions including **sum**, **avg**, **max**, **min**, and **count** are supported. Unsupported scenarios: • **aggregate** contains operations, for example, sum(a+b). • **aggregate** operations, for example, sum(a)+min(b).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Precautions

- If external data source is Hive, query operation cannot be performed on foreign tables created by Spark.
- Only MySQL and MPPDB data sources are supported.

1.6.30 Spark2x

1.6.30.1 Spark2x Basic Principles

Description

Spark is a memory-based distributed computing framework. In iterative computation scenarios, the computing capability of Spark is 10 to 100 times higher than MapReduce, because data is stored in memory when being processed. Spark can use HDFS as the underlying storage system, enabling users to quickly switch to Spark from MapReduce. Spark provides one-stop data analysis capabilities, such as the streaming processing in small batches, offline batch processing, SQL query, and data mining. Users can seamlessly use these functions in a same application. For details about the new open-source features of Spark2x, see **Spark2x Open Source New Features**.

Features of Spark are as follows:

- Improves the data processing capability through distributed memory computing and directed acyclic graph (DAG) execution engine. The delivered performance is 10 to 100 times higher than that of MapReduce.
- Supports multiple development languages (Scala/Java/Python) and dozens of highly abstract operators to facilitate the construction of distributed data processing applications.
- Builds data processing stacks using **SQL**, **Streaming**, MLlib, and GraphX to provide one-stop data processing capabilities.
- Fits into the Hadoop ecosystem, allowing Spark applications to run on Standalone, Mesos, or Yarn, enabling access of multiple data sources such as HDFS, HBase, and Hive, and supporting smooth migration of the MapReduce application to Spark.

Architecture

Figure 1-112 describes the Spark architecture and **Table 1-26** lists the Spark modules.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-112 Spark architecture

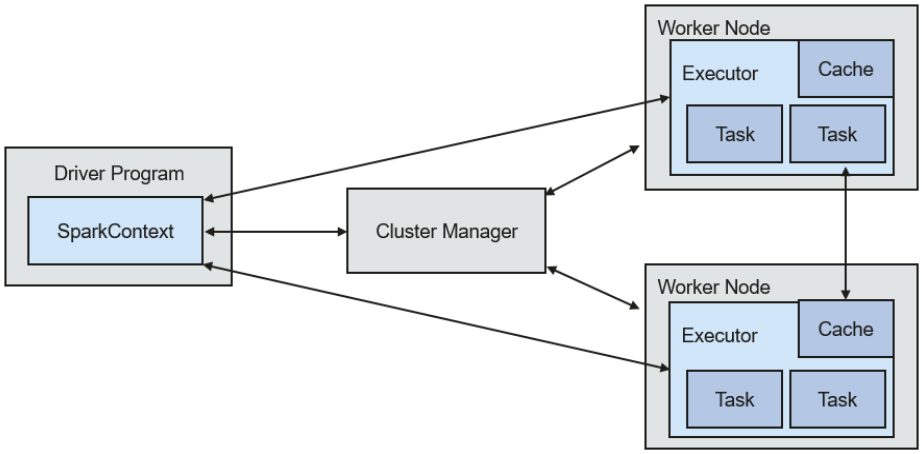


Table 1-26 Basic concepts

Module	Description
Cluster Manager	Cluster manager manages resources in the cluster. Spark supports multiple cluster managers, including Mesos, Yarn, and the Standalone cluster manager that is delivered with Spark. By default, Spark clusters adopt the Yarn cluster manager.
Application	Spark application. It consists of one Driver Program and multiple executors.
Deploy Mode	Deployment in cluster or client mode. In cluster mode, the driver runs on a node inside the cluster. In client mode, the driver runs on the client (outside the cluster).
Driver Program	The main process of the Spark application. It runs the **main()** function of an application and creates SparkContext. It is used for parsing applications, generating stages, and scheduling tasks to executors. Usually, SparkContext represents Driver Program.
Executor	A process started on a Work Node. It is used to execute tasks, and manage and process the data used in applications. A Spark application usually contains multiple executors. Each executor receives commands from the driver and executes one or multiple tasks.
Worker Node	A node that starts and manages executors and resources in a cluster.
Job	A job consists of multiple concurrent tasks. One action operator (for example, a collect operator) maps to one job.
Stage	Each job consists of multiple stages. Each stage is a task set, which is separated by Directed Acyclic Graph (DAG).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

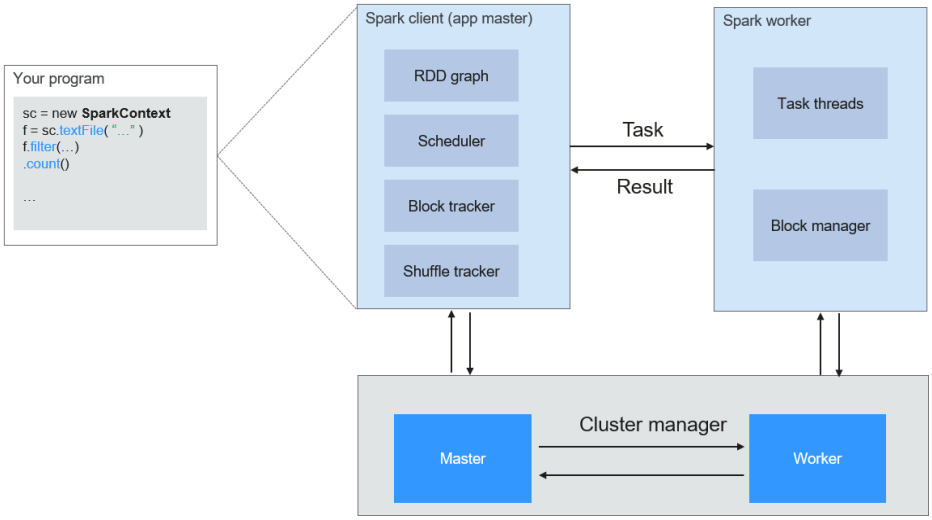
Module	Description
Task	A task carries the computation unit of the service logics. It is the minimum working unit that can be executed on the Spark platform. An application can be divided into multiple tasks based on the execution plan and computation amount.

Spark Principle

Figure 1-113 describes the application running architecture of Spark.

1. An application is running in the cluster as a collection of processes. Driver coordinates the running of the application.
2. To run an application, Driver connects to the cluster manager (such as Standalone, Mesos, and Yarn) to apply for the executor resources, and start ExecutorBackend. The cluster manager schedules resources between different applications. Driver schedules DAGs, divides stages, and generates tasks for the application at the same time.
3. Then, Spark sends the code of the application (the code transferred to SparkContext, which is defined by JAR or Python) to an executor.
4. After all tasks are finished, the running of the user application is stopped.

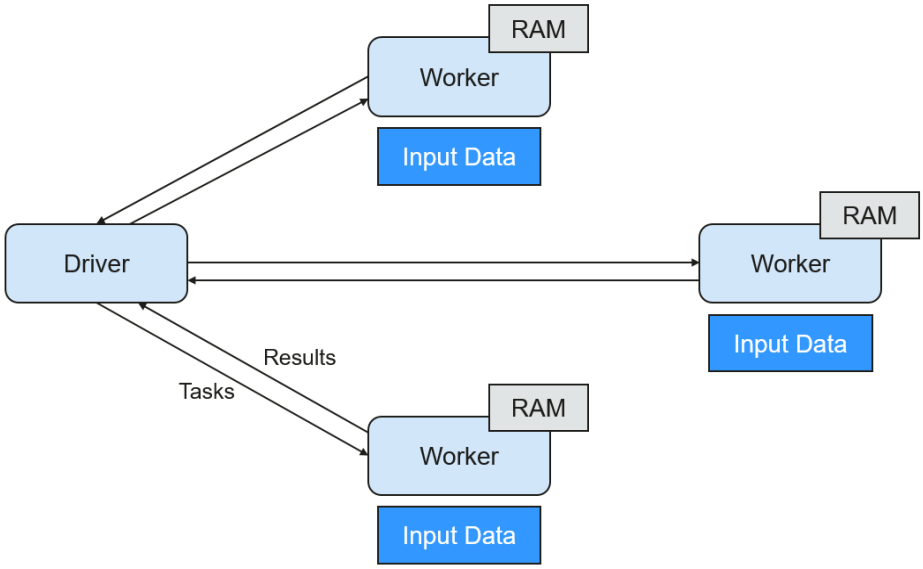
Figure 1-113 Spark application running architecture



Spark uses Master and Worker modes, as shown in Figure 1-114. A user submits an application on the Spark client, and then the scheduler divides a job into multiple tasks and sends the tasks to each Worker for execution. Each Worker reports the computation results to Driver (Master), and then the Driver aggregates and returns the results to the client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-114 Spark Master-Worker mode



Note the following about the architecture:

- Applications are isolated from each other.

  Each application has an independent executor process, and each executor starts multiple threads to execute tasks in parallel. Each driver schedules its own tasks, and different application tasks run on different JVMs, that is, different executors.
- Different Spark applications do not share data, unless data is stored in the external storage system such as HDFS.
- You are advised to deploy the Driver program in a location that is close to the Worker node because the Driver program schedules tasks in the cluster. For example, deploy the Driver program on the network where the Worker node is located.

Spark on YARN can be deployed in two modes:

- In Yarn-cluster mode, the Spark driver runs inside an ApplicationMaster process which is managed by Yarn in the cluster. After the ApplicationMaster is started, the client can exit without interrupting service running.
- In Yarn-client mode, Driver runs in the client process, and the ApplicationMaster process is used only to apply for requesting resources from Yarn.

Spark Streaming Principle

Spark Streaming is a real-time computing framework built on the Spark, which expands the capability for processing massive streaming data. Spark supports two data processing approaches: Direct Streaming and Receiver.

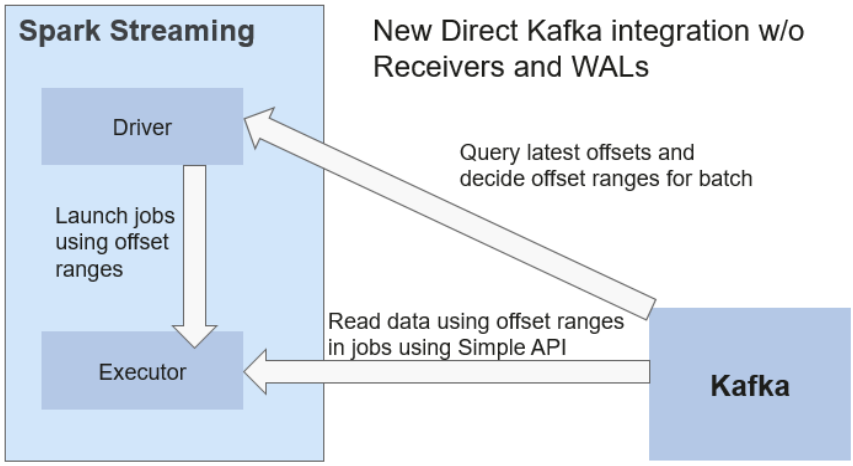
Direct Streaming computing process

In Direct Streaming approach, Direct API is used to process data. Take Kafka Direct API as an example. Direct API provides offset location that each batch range will read from, which is much simpler than starting a receiver to continuously receive

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

data from Kafka and written data to write-ahead logs (WALs). Then, each batch job is running and the corresponding offset data is ready in Kafka. These offset information can be securely stored in the checkpoint file and read by applications that failed to start.

Figure 1-115 Data transmission through Direct Kafka API



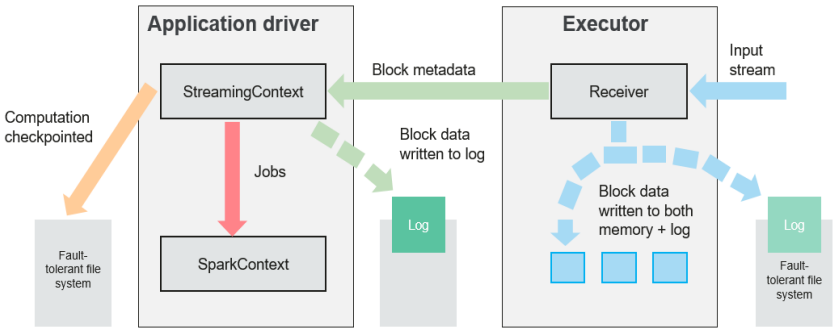
After the failure, Spark Streaming can read data from Kafka again and process the data segment. The processing result is the same no matter Spark Streaming fails or not, because the semantic is processed only once.

Direct API does not need to use the WAL and Receivers, and ensures that each Kafka record is received only once, which is more efficient. In this way, the Spark Streaming and Kafka can be well integrated, making streaming channels be featured with high fault-tolerance, high efficiency, and ease-of-use. Therefore, you are advised to use Direct Streaming to process data.

Receiver computing process

When a Spark Streaming application starts (that is, when the driver starts), the related StreamingContext (the basis of all streaming functions) uses SparkContext to start the receiver to become a long-term running task. These receivers receive and save streaming data to the Spark memory for processing. **Figure 1-116** shows the data transfer lifecycle.

Figure 1-116 Data transfer lifecycle



1. Receive data (blue arrow).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Receiver divides a data stream into a series of blocks and stores them in the executor memory. In addition, after WAL is enabled, it writes data to the WAL of the fault-tolerant file system.

2. Notify the driver (green arrow).

   The metadata in the received block is sent to StreamingContext in the driver. The metadata includes:

   - Block reference ID used to locate the data position in the Executor memory.
   - Block data offset information in logs (if the WAL function is enabled).

3. Process data (red arrow).

   For each batch of data, StreamingContext uses block information to generate resilient distributed datasets (RDDs) and jobs. StreamingContext executes jobs by running tasks to process blocks in the executor memory.

4. Periodically set checkpoints (orange arrows).
5. For fault tolerance, StreamingContext periodically sets checkpoints and saves them to external file systems.

Fault Tolerance

Spark and its RDD allow seamless processing of failures of any Worker node in the cluster. Spark Streaming is built on top of Spark. Therefore, the Worker node of Spark Streaming also has the same fault tolerance capability. However, Spark Streaming needs to run properly in case of long-time running. Therefore, Spark must be able to recover from faults through the driver process (main process that coordinates all Workers). This poses challenges to the Spark driver fault-tolerance because the Spark driver may be any user application implemented in any computation mode. However, Spark Streaming has internal computation architecture. That is, it periodically executes the same Spark computation in each batch data. Such architecture allows it to periodically store checkpoints to reliable storage space and recover them upon the restart of Driver.

For source data such as files, the Driver recovery mechanism can ensure zero data loss because all data is stored in a fault-tolerant file system such as HDFS. However, for other data sources such as Kafka and Flume, some received data is cached only in memory and may be lost before being processed. This is caused by the distribution operation mode of Spark applications. When the driver process fails, all executors running in the Cluster Manager, together with all data in the memory, are terminated. To avoid such data loss, the WAL function is added to Spark Streaming.

WAL is often used in databases and file systems to ensure persistence of any data operation. That is, first record an operation to a persistent log and perform this operation on data. If the operation fails, the system is recovered by reading the log and re-applying the preset operation. The following describes how to use WAL to ensure persistence of received data:

Receiver is used to receive data from data sources such as Kafka. As a long-time running task in Executor, Receiver receives data, and also confirms received data if supported by data sources. Received data is stored in the Executor memory, and Driver delivers a task to Executor for processing.

After WAL is enabled, all received data is stored to log files in the fault-tolerant file system. Therefore, the received data does not lose even if Spark Streaming

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

fails. Besides, receiver checks correctness of received data only after the data is pre-written into logs. Data that is cached but not stored can be sent again by data sources after the driver restarts. These two mechanisms ensure zero data loss. That is, all data is recovered from logs or re-sent by data sources.

To enable the WAL function, perform the following operations:

- Set **streamingContext.checkpoint** (path-to-directory) to configure the checkpoint directory, which is an HDFS file path used to store streaming checkpoints and WALs.
- Set **spark.streaming.receiver.writeAheadLog.enable** of SparkConf to **true** (the default value is **false**).

After WAL is enabled, all receivers have the advantage of recovering from reliable received data. You are advised to disable the multi-replica mechanism because the fault-tolerant file system of WAL may also replicate the data.

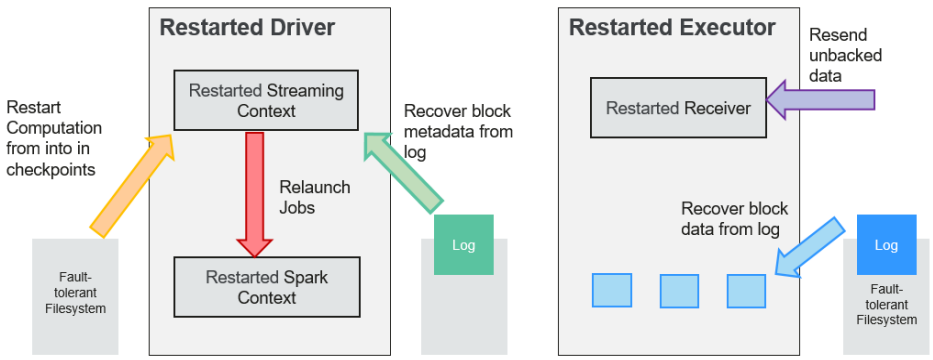
NOTE

The data receiving throughput is lowered after WAL is enabled. All data is written into the fault-tolerant file system. As a result, the write throughput of the file system and the network bandwidth for data replication may become the potential bottleneck. To solve this problem, create more receivers to increase the data receiving parallelism or use better hardware to improve the throughput of the fault-tolerant file system.

Recovery Process

When a failed driver is restarted, restart it as follows:

Figure 1-117 Computing recovery process



1. Recover computing. (Orange arrow)

   Use checkpoint information to restart Driver, reconstruct SparkContext and restart Receiver.

2. Recover metadata block. (Green arrow)

   This operation ensures that all necessary metadata blocks are recovered to continue the subsequent computing recovery.

3. Relaunch unfinished jobs. (Red arrow)

   Recovered metadata is used to generate RDDs and corresponding jobs for interrupted batch processing due to failures.

4. Read block data saved in logs. (Blue arrow)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Block data is directly read from WALs during execution of the preceding jobs, and therefore all essential data reliably stored in logs is recovered.

5. Resend unconfirmed data. (Purple arrow)

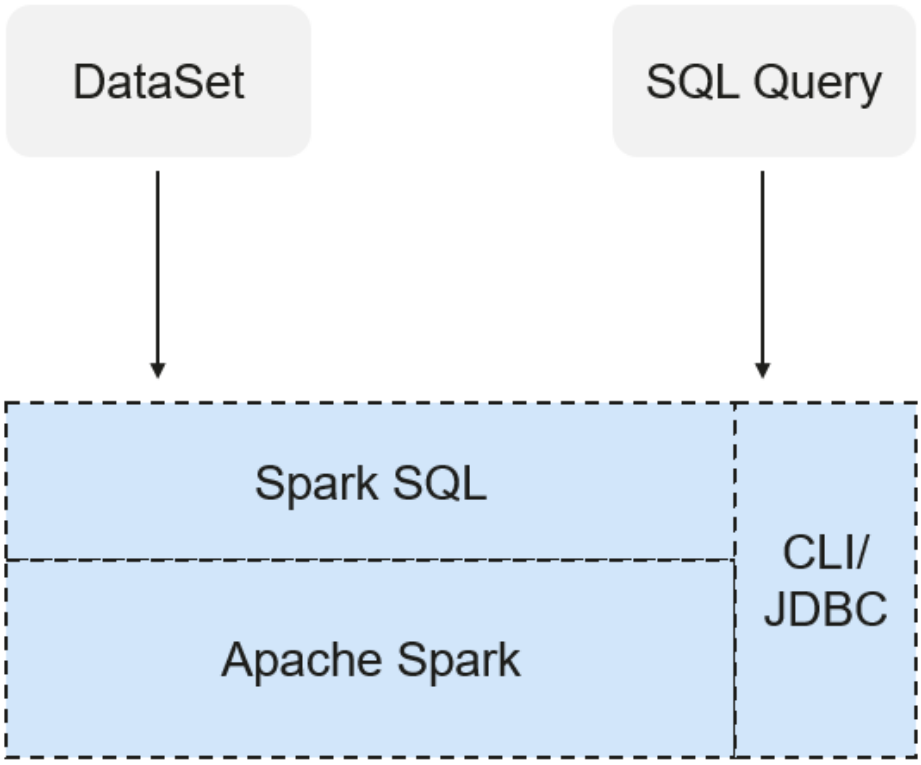
   Data that is cached but not stored to logs upon failures is re-sent by data sources, because the receiver does not confirm the data.

Therefore, by using WALs and reliable Receiver, Spark Streaming can avoid input data loss caused by Driver failures.

SparkSQL and DataSet Principle

SparkSQL

Figure 1-118 SparkSQL and DataSet



Spark SQL is a module for processing structured data. In Spark applications, SQL statements or DataSet APIs can be seamlessly used for querying structured data.

Spark SQL and DataSet also provide a universal method for accessing multiple data sources such as Hive, CSV, Parquet, ORC, JSON, and JDBC. These data sources also allow data interaction. Spark SQL reuses the Hive frontend processing logic and metadata processing module. With the Spark SQL, you can directly query existing Hive data.

In addition, Spark SQL also provides API, CLI, and JDBC APIs, allowing diverse accesses to the client.

Spark SQL Native DDL/DML

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In Spark 1.5, lots of Data Definition Language (DDL)/Data Manipulation Language (DML) commands are pushed down to and run on the Hive, causing coupling with the Hive and inflexibility such as unexpected error reports and results.

Spark2x realizes command localization and replaces the Hive with Spark SQL Native DDL/DML to run DDL/DML commands. Additionally, the decoupling from the Hive is realized and commands can be customized.

DataSet

A DataSet is a strongly typed collection of domain-specific objects that can be transformed in parallel using functional or relational operations. Each Dataset also has an untyped view called a DataFrame, which is a Dataset of Row.

The DataFrame is a structured and distributed dataset consisting of multiple columns. The DataFrame is equal to a table in the relationship database or the DataFrame in the R/Python. The DataFrame is the most basic concept in the Spark SQL, which can be created by using multiple methods, such as the structured dataset, Hive table, external database or RDD.

Operations available on DataSets are divided into transformations and actions.

- A transformation operation can generate a new DataSet,
  for example, **map**, **filter**, **select**, and **aggregate (groupBy)**.
- An action operation can trigger computation and return results,
  for example, **count**, **show**, or write data to the file system.

You can use either of the following methods to create a DataSet:

- The most common way is by pointing Spark to some files on storage systems, using the **read** function available on a SparkSession.
  ```
  val people = spark.read.parquet("...").as[Person]  // Scala
  DataSet<Person> people = spark.read().parquet("...").as(Encoders.bean(Person.class));//Java
  ```
- You can also create a DataSet using the transformation operation available on an existing one. For example, apply the map operation on an existing DataSet to create a DataSet:
  ```
  val names = people.map(_.name) // In Scala: names is Dataset.
  Dataset<String> names = people.map((Person p) -> p.name, Encoders.STRING)); // Java
  ```

CLI and JDBCServer

In addition to programming APIs, Spark SQL also provides the CLI/JDBC APIs.

- Both **spark-shell** and **spark-sql** scripts can provide the CLI for debugging.
- JDBCServer provides JDBC APIs. External systems can directly send JDBC requests to calculate and parse structured data.

SparkSession Principle

SparkSession is a unified API in Spark2x and can be regarded as a unified entry for reading data. SparkSession provides a single entry point to perform many operations that were previously scattered across multiple classes, and also provides accessor methods to these older classes to maximize compatibility.

A SparkSession can be created using a builder pattern. The builder will automatically reuse the existing SparkSession if there is a SparkSession; or create a SparkSession if it does not exist. During I/O transactions, the configuration item settings in the builder are automatically synchronized to Spark and Hadoop.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
import org.apache.spark.sql.SparkSession
val sparkSession = SparkSession.builder
  .master("local")
  .appName("my-spark-app")
  .config("spark.some.config.option", "config-value")
  .getOrCreate()
```

- SparkSession can be used to execute SQL queries on data and return results as DataFrame.
  ```
  sparkSession.sql("select * from person").show
  ```
- SparkSession can be used to set configuration items during running. These configuration items can be replaced with variables in SQL statements.
  ```
  sparkSession.conf.set("spark.some.config", "abcd")
  sparkSession.conf.get("spark.some.config")
  sparkSession.sql("select ${spark.some.config}")
  ```
- SparkSession also includes a "catalog" method that contains methods to work with Metastore (data catalog). After this method is used, a dataset is returned, which can be run using the same Dataset API.
  ```
  val tables = sparkSession.catalog.listTables()
  val columns = sparkSession.catalog.listColumns("myTable")
  ```
- Underlying SparkContext can be accessed by SparkContext API of SparkSession.
  ```
  val sparkContext = sparkSession.sparkContext
  ```

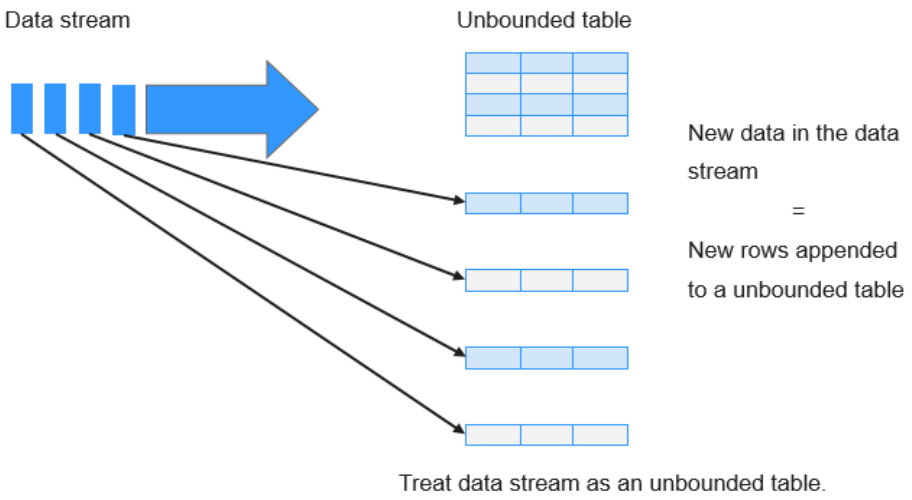
Structured Streaming Principle

Structured Streaming is a stream processing engine built on the Spark SQL engine. You can use the Dataset/DataFrame API in Scala, Java, Python, or R to express streaming aggregations, event-time windows, and stream-stream joins. If streaming data is incrementally and continuously produced, Spark SQL will continue to process the data and synchronize the result to the result set. In addition, the system ensures end-to-end exactly-once fault-tolerance guarantees through checkpoints and WALs.

The core of Structured Streaming is to take streaming data as an incremental database table. Similar to the data block processing model, the streaming data processing model applies query operations on a static database table to streaming computing, and Spark uses standard SQL statements for query, to obtain data from the incremental and unbounded table.

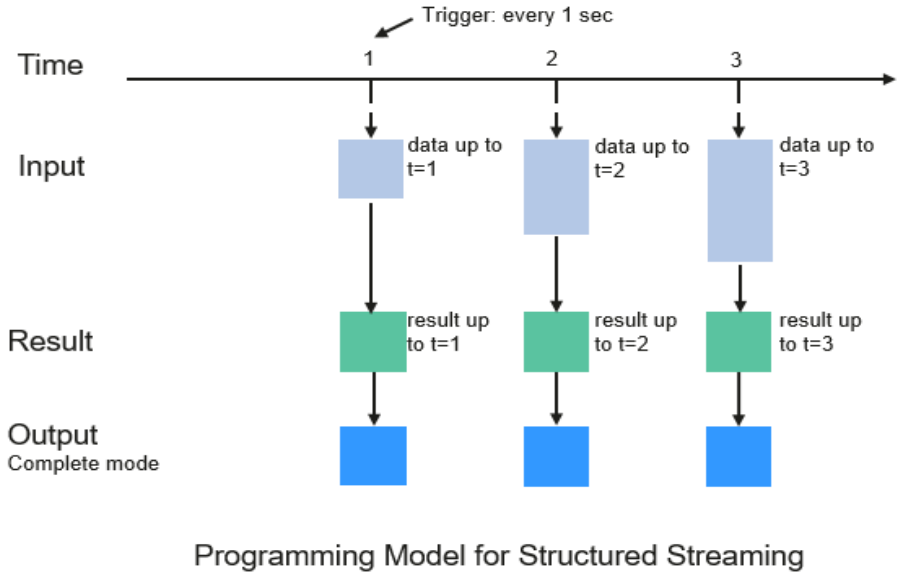
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-119 Unbounded table of Structured Streaming



Each query operation will generate a result table. At each trigger interval, updated data will be synchronized to the result table. Whenever the result table is updated, the updated result will be written into an external storage system.

Figure 1-120 Structured Streaming data processing model



Storage modes of Structured Streaming at the output phase are as follows:

- Complete Mode: The updated result sets are written into the external storage system. The write operation is performed by a connector of the external storage system.
- Append Mode: If an interval is triggered, only added data in the result table will be written into an external system. This is applicable only to the queries where existing rows in the result table are not expected to change.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Update Mode: If an interval is triggered, only updated data in the result table will be written into an external system, which is the difference between the Complete Mode and Update Mode.

Concepts

- **RDD**

  Resilient Distributed Dataset (RDD) is a core concept of Spark. It indicates a read-only and partitioned distributed dataset. Partial or all data of this dataset can be cached in the memory and reused between computations.

  RDD Creation

  - An RDD can be created from the input of HDFS or other storage systems that are compatible with Hadoop.
  - A new RDD can be converted from a parent RDD.
  - An RDD can be converted from a collection of datasets through encoding.

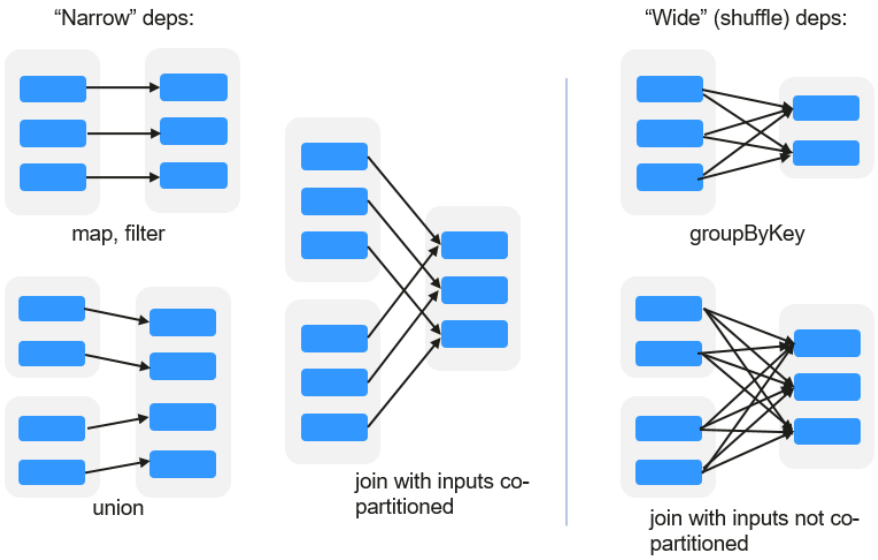
  RDD Storage

  - You can select different storage levels to store an RDD for reuse. (There are 11 storage levels to store an RDD.)
  - By default, the RDD is stored in the memory. When the memory is insufficient, the RDD overflows to the disk.

- **RDD Dependency**

  The RDD dependency includes the narrow dependency and wide dependency.

  Figure 1-121 RDD dependency

  - **Narrow dependency**: Each partition of the parent RDD is used by at most one partition of the child RDD.
  - **Wide dependency**: Partitions of the child RDD depend on all partitions of the parent RDD.

  The narrow dependency facilitates the optimization. Logically, each RDD operator is a fork/join (the join is not the join operator mentioned above but the barrier used to synchronize multiple concurrent tasks); fork the RDD to

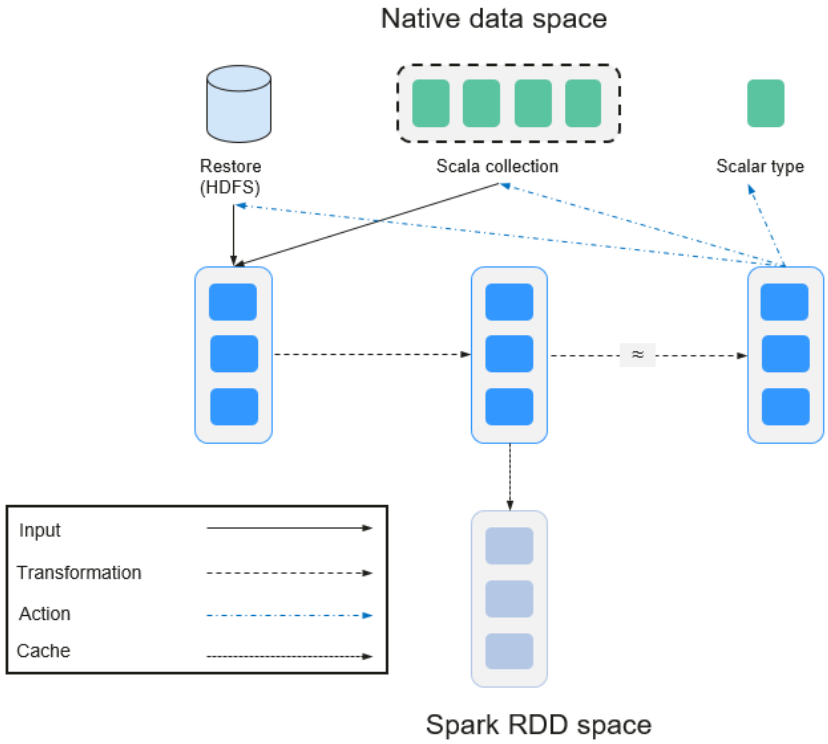
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

each partition, and then perform the computation. After the computation, join the results, and then perform the fork/join operation on the next RDD operator. It is uneconomical to directly translate the RDD into physical implementation. The first is that every RDD (even intermediate result) needs to be physicalized into memory or storage, which is time-consuming and occupies much space. The second is that as a global barrier, the join operation is very expensive and the entire join process will be slowed down by the slowest node. If the partitions of the child RDD narrowly depend on that of the parent RDD, the two fork/join processes can be combined to implement classic fusion optimization. If the relationship in the continuous operator sequence is narrow dependency, multiple fork/join processes can be combined to reduce a large number of global barriers and eliminate the physicalization of many RDD intermediate results, which greatly improves the performance. This is called pipeline optimization in Spark.

- **Transformation and Action (RDD Operations)**

  Operations on RDD include transformation (the return value is an RDD) and action (the return value is not an RDD). **Figure 1-122** shows the RDD operation process. The transformation is lazy, which indicates that the transformation from one RDD to another RDD is not immediately executed. Spark only records the transformation but does not execute it immediately. The real computation is started only when the action is started. The action returns results or writes the RDD data into the storage system. The action is the driving force for Spark to start the computation.

  Figure 1-122 RDD operation



The data and operation model of RDD are quite different from those of Scala.

```
val file = sc.textFile("hdfs://...")
val errors = file.filter(_.contains("ERROR"))
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
errors.cache()
errors.count()
```

a. The textFile operator reads log files from the HDFS and returns files (as an RDD).
b. The filter operator filters rows with **ERROR** and assigns them to errors (a new RDD). The filter operator is a transformation.
c. The cache operator caches errors for future use.
d. The count operator returns the number of rows of errors. The count operator is an action.

Transformation includes the following types:

- The RDD elements are regarded as simple elements.

  The input and output has the one-to-one relationship, and the partition structure of the result RDD remains unchanged, for example, map.

  The input and output has the one-to-many relationship, and the partition structure of the result RDD remains unchanged, for example, flatMap (one element becomes a sequence containing multiple elements after map and then flattens to multiple elements).

  The input and output has the one-to-one relationship, but the partition structure of the result RDD changes, for example, union (two RDDs integrates to one RDD, and the number of partitions becomes the sum of the number of partitions of two RDDs) and coalesce (partitions are reduced).

  Operators of some elements are selected from the input, such as filter, distinct (duplicate elements are deleted), subtract (elements only exist in this RDD are retained), and sample (samples are taken).

- The RDD elements are regarded as key-value pairs.

  Perform the one-to-one calculation on the single RDD, such as mapValues (the partition mode of the source RDD is retained, which is different from map).

  Sort the single RDD, such as sort and partitionBy (partitioning with consistency, which is important to the local optimization).

  Restructure and reduce the single RDD based on key, such as groupByKey and reduceByKey.

  Join and restructure two RDDs based on the key, such as join and cogroup.

  NOTE

  The later three operations involving sorting are called shuffle operations.

Action includes the following types:

- Generate scalar configuration items, such as **count** (the number of elements in the returned RDD), **reduce**, **fold/aggregate** (the number of scalar configuration items that are returned), and **take** (the number of elements before the return).
- Generate the Scala collection, such as **collect** (import all elements in the RDD to the Scala collection) and **lookup** (look up all values corresponding to the key).
- Write data to the storage, such as **saveAsTextFile** (which corresponds to the preceding **textFile**).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

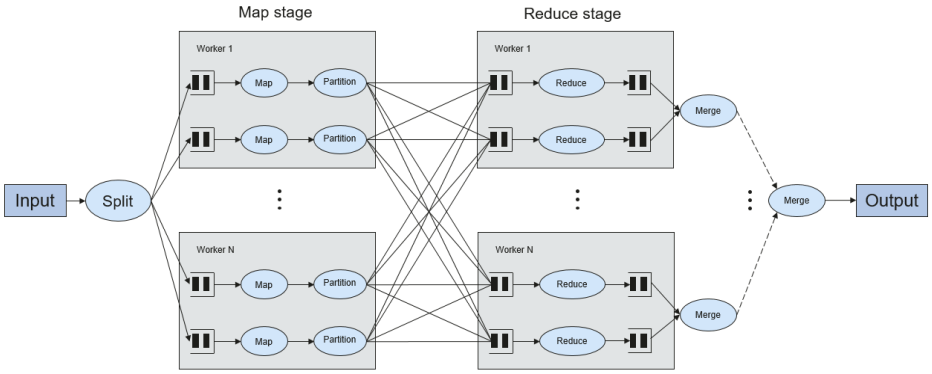
  - Check points, such as the **checkpoint** operator. When Lineage is quite long (which occurs frequently in graphics computation), it takes a long period of time to execute the whole sequence again when a fault occurs. In this case, checkpoint is used as the check point to write the current data to stable storage.

- **Shuffle**

  Shuffle is a specific phase in the MapReduce framework, which is located between the Map phase and the Reduce phase. If the output results of Map are to be used by Reduce, the output results must be hashed based on a key and distributed to each Reducer. This process is called Shuffle. Shuffle involves the read and write of the disk and the transmission of the network, so that the performance of Shuffle directly affects the operation efficiency of the entire program.

  The figure below shows the entire process of the MapReduce algorithm.

  Figure 1-123 Algorithm process



  Shuffle is a bridge connecting data. The following describes the implementation of shuffle in Spark.

  Shuffle divides a job of Spark into multiple stages. The former stages contain one or more ShuffleMapTasks, and the last stage contains one or more ResultTasks.

- **Spark Application Structure**

  The Spark application structure includes the initialized SparkContext and the main program.

  - Initialized SparkContext: constructs the operating environment of the Spark Application.

    Constructs the SparkContext object. The following is an example:

    ```
    new SparkContext(master, appName, [SparkHome], [jars])
    ```

    Parameter description:

    master: indicates the link string. The link modes include local, Yarn-cluster, and Yarn-client.

    appName: indicates the application name.

    SparkHome: indicates the directory where Spark is installed in the cluster.

    jars: indicates the code and dependency package of an application.

  - Main program: processes data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Spark Shell Commands**

  The basic Spark shell commands support the submission of Spark applications. The Spark shell commands are as follows:

  ```
  ./bin/spark-submit \
    --class <main-class> \
    --master <master-url> \
    ... # other options
    <application-jar> \
    [application-arguments]
  ```

  Parameter description:

  --class: indicates the name of the class of a Spark application.

  --master: indicates the master to which the Spark application links, such as Yarn-client and Yarn-cluster.

  application-jar: indicates the path of the JAR file of the Spark application.

  application-arguments: indicates the parameter required to submit the Spark application. This parameter can be left blank.

- **Spark JobHistory Server**

  The Spark web UI is used to monitor the details in each phase of the Spark framework of a running or historical Spark job and provide the log display, which helps users to develop, configure, and optimize the job in more fine-grained units.

1.6.30.2 Spark2x Multi-active Instance

Background

Based on existing JDBCServers in the community, multi-active-instance HA is used to achieve the high availability. In this mode, multiple JDBCServers coexist in the cluster and the client can randomly connect any JDBCServer to perform service operations. When one or multiple JDBCServers stop working, a client can connect to another normal JDBCServer.

Compared with active/standby HA, multi-active instance HA eliminates the following restrictions:

- In active/standby HA, when the active/standby switchover occurs, the unavailable period cannot be controlled by JDBCServer, but determined by Yarn service resources.
- In Spark, the Thrift JDBC similar to HiveServer2 provides services and users access services through Beeline and JDBC API. Therefore, the processing capability of the JDBCServer cluster depends on the single-point capability of the primary server, and the scalability is insufficient.

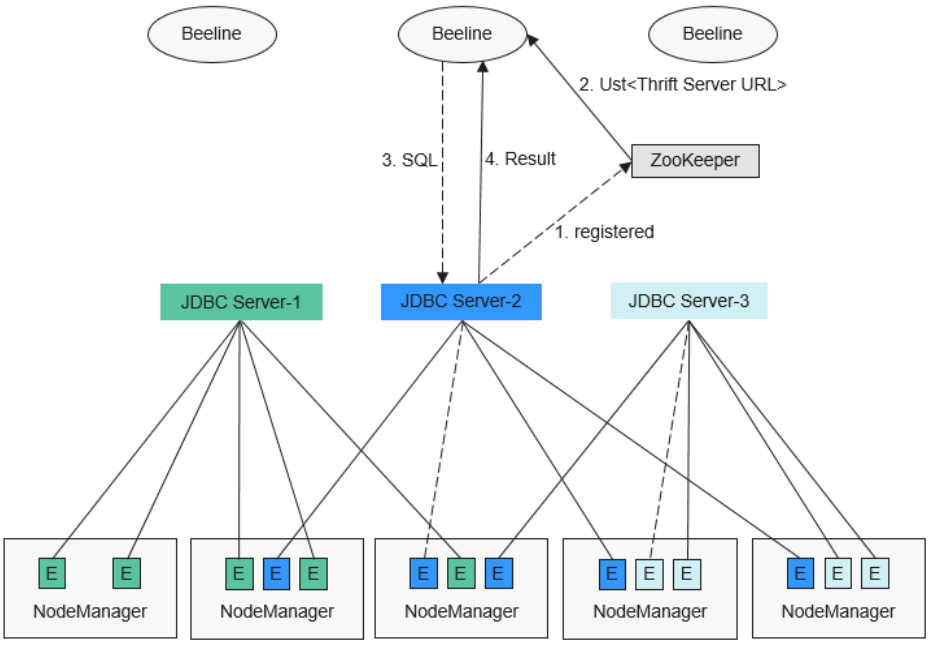
Multi-active instance HA not only prevents service interruption caused by switchover, but also enables cluster scale-out to secure high concurrency.

Implementation

The following figure shows the basic principle of multi-active instance HA of Spark JDBCServer.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-124 Spark JDBCServer HA



1. After JDBCServer is started, it registers with ZooKeeper by writing node information in a specified directory. Node information includes the JDBCServer instance IP, port number, version, and serial number (information of different nodes is separated by commas).

   An example is provided as follows:

   ```
   [serverUri=192.168.169.84:22550
   ;version=8.1.0.1;sequence=0000001244,serverUri=192.168.195.232:22550 ;version=8.1.0.1;sequence=000
   0001242,serverUri=192.168.81.37:22550;version=8.1.0.1;sequence=0000001243]
   ```

2. To connect to JDBCServer, the client must specify the namespace, which is the directory of JDBCServer instances in ZooKeeper. During the connection, a JDBCServer instance is randomly selected from the specified namespace. For details about URL, see **URL Connection**.
3. After the connection succeeds, the client sends SQL statements to JDBCServer.
4. JDBCServer executes received SQL statements and sends results back to the client.

In multi-active instance HA mode, all JDBCServer instances are independent and equivalent. When one instance is interrupted during upgrade, other JDBCServer instances can accept the connection request from the client.

Following rules must be followed in the multi-active instance HA of Spark JDBCServer:

- If a JDBCServer instance exits abnormally, no other instance will take over the sessions and services running on this abnormal instance.
- When the JDBCServer process is stopped, corresponding nodes are deleted from ZooKeeper.
- The client randomly selects the server, which may result in uneven session allocation, and finally result in imbalance of instance load.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After the instance enters the maintenance mode (in which no new connection request from the client is accepted), services still running on the instance may fail when the decommissioning times out.

URL Connection

Multi-active instance mode

In multi-active instance mode, the client reads content from the ZooKeeper node and connects to JDBCServer. The connection strings are as follows:

- Security mode:
  - If Kinit authentication is enabled, the JDBCURL is as follows:
    jdbc:hive2://
    <zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>
    /;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.*<System domain name>***@***<System domain name>***;**

    NOTE

    - **<zkNode_IP>:<zkNode_Port>** indicates the ZooKeeper URL. Use commas (,) to separate multiple URLs.

      For example,
      192.168.81.37:2181,192.168.195.232:2181,192.168.169.84:2181.
    - **sparkthriftserver2x** indicates the directory in ZooKeeper, where a random JDBCServer instance is connected to the client.

    For example, when you use Beeline client for connection in security mode, run the following command:

    sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://***<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.***<System domain name>***@***<System domain name>***;"**
  - If Keytab authentication is enabled, the JDBCURL is as follows:
    jdbc:hive2:*//*
    *<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.***<System domain name>***@***<System domain name>***;user.principal=***<principal_name>***;user.keytab=***<path_to_keytab>**

    <principal_name> indicates the principal of Kerberos user, for example, **test@***<System domain name>*. *<path_to_keytab>* indicates the Keytab file path corresponding to *<principal_name>*, for example, **/opt/auth/test/user.keytab**.
- Common mode:
  jdbc:hive2://
  <zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>
  /;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=**sparkthriftserver2x**;

  For example, when you use Beeline client for connection in common mode, run the following command:

  sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://***<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;"**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Non-multi-active instance mode

In non-multi-active instance mode, a client connects to a specified JDBCServer node. Compared with multi-active instance mode, the connection string in non-multi-active instance mode does not contain **serviceDiscoveryMode** and **zooKeeperNamespace** parameters about ZooKeeper.

For example, when you use Beeline client to connect JDBCServer in non-multi-active instance mode, run the following command:

sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://**<*server_IP*>:<*server_Port*>**/;user.principal=spark2x/hadoop.**<*System domain name*>@<*System domain name*>**;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.**<*System domain name*>@<*System domain name*>**;"**

NOTE

- **<server_IP>:<server_Port>** indicates the URL of the specified JDBCServer node.
- **CLIENT_HOME** indicates the client path.

Except the connection method, operations of JDBCServer API in multi-active instance mode and non-multi-active instance mode are the same.

1.6.30.3 Spark2x Multi-tenant

Background

In the JDBCServer multi-active instance mode, JDBCServer implements the Yarn-client mode but only one Yarn resource queue is available. To solve the resource limitation problem, the multi-tenant mode is introduced.

In multi-tenant mode, JDBCServers are bound with tenants. Each tenant corresponds to one or more JDBCServers, and a JDBCServer provides services for only one tenant. Different tenants can be configured with different Yarn queues to implement resource isolation. In addition, JDBCServer can be dynamically started as required to avoid resource waste.

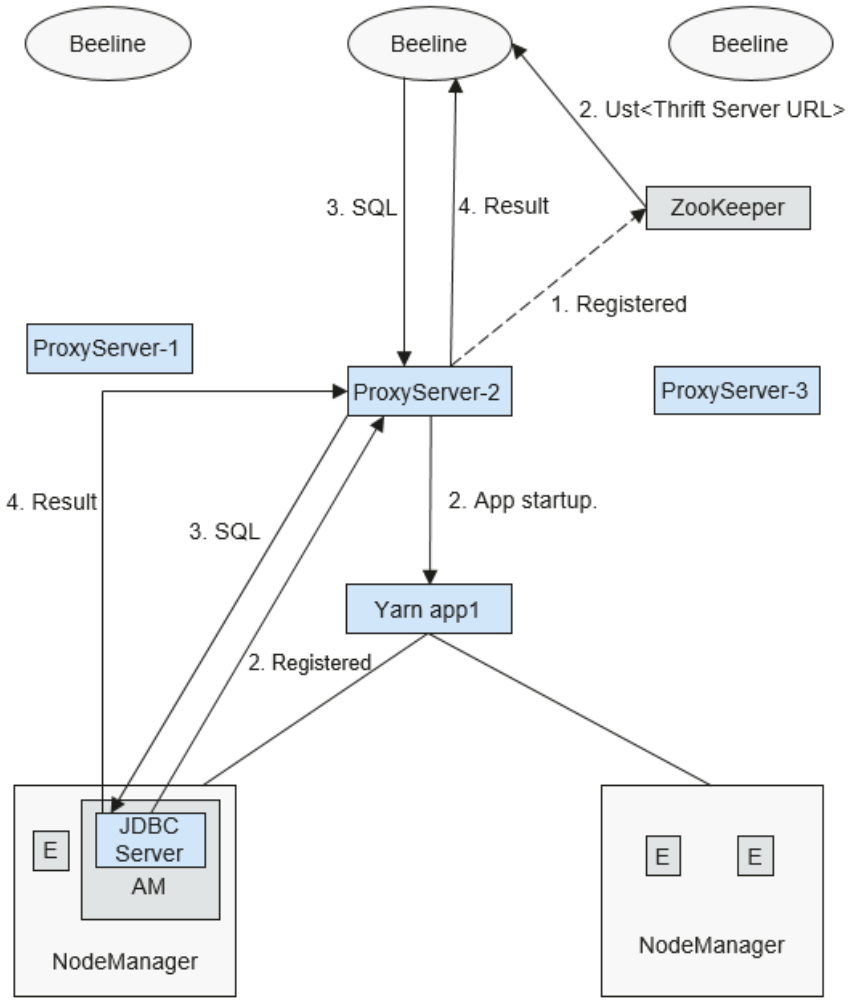
MRS 3.5.0-LTS and later versions support two complex balancing policies in the multi-tenant mode of JDBCServer: **DRIVER_METRICS** and **ORIGIN**. The **DRIVER_METRICS** algorithm is used by default.

Implementation

Figure 1-125 shows the HA solution of the multi-tenant mode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-125 Multi-tenant mode of Spark JDBCServer



1. When ProxyServer is started, it registers with ZooKeeper by writing node information in a specified directory. Node information includes the instance IP, port number, version, and serial number (information of different nodes is separated by commas).

   NOTE

   In multi-tenant mode, the JDBCServer instance on MRS page indicates ProxyServer, the JDBCServer agent.

   An example is provided as follows:

   ```
   serverUri=192.168.169.84:22550
   ;version=8.1.0.1;sequence=0000001244,serverUri=192.168.195.232:22550
   ;version=8.1.0.1;sequence=0000001242,serverUri=192.168.81.37:22550
   ;version=8.1.0.1;sequence=0000001243,
   ```

2. To connect to ProxyServer, the client must specify a namespace, which is the directory of the ProxyServer instance that you want to access in ZooKeeper. During the connection, the system performs the modulo operation on the hash value of the tenant name and the number of namespace instances in ZooKeeper to obtain the connected instance. For details about the URL, see **URL Connection**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. After the client successfully connects to ProxyServer, ProxyServer checks whether the JDBCServer of a tenant exists. If yes, Beeline connects the JDBCServer. If no, a new JDBCServer is started in Yarn-cluster mode. After the startup of JDBCServer, ProxyServer obtains the IP address of the JDBCServer and establishes the connection between Beeline and JDBCServer.
4. The client sends SQL statements to ProxyServer, which then forwards statements to the connected JDBCServer. JDBCServer returns the results to ProxyServer, which then returns the results to the client.

In multi-tenant HA mode, all ProxyServer instances are independent and equivalent. If one instance is interrupted during upgrade, other instances can accept the connection request from the client.

URL Connection

Multi-tenant mode

In multi-tenant mode, the client reads content from the ZooKeeper node and connects to ProxyServer. The connection strings are as follows:

- Security mode:
  - If Kinit authentication is enabled, the client URL is as follows:
    jdbc:hive2:// *<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>*/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.*<System domain name>*@*<System domain name>*;

    NOTE

    - **<zkNode_IP>:<zkNode_Port>** indicates the ZooKeeper URL. Use commas (,) to separate multiple URLs.

      For example, **192.168.81.37:2181,192.168.195.232:2181,192.168.169.84:2181**.
    - **sparkthriftserver2x** indicates the ZooKeeper directory, where a random JDBCServer instance is connected to the client.

    For example, when you use Beeline client for connection in security mode, run the following command:

    sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://** *<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>*/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.*<System domain name>*@*<System domain name>*;"
  - If Keytab authentication is enabled, the URL is as follows:
    jdbc:hive2:// *<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>*/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.*<System domain name>*@*<System domain name>*;user.principal=*<principal_name>*;user.keytab=*<path_to_keytab>*

    <principal_name> indicates the principal of Kerberos user, for example, **test@***<System domain name>*. *<path_to_keytab>* indicates the Keytab file path corresponding to *<principal_name>*, for example, **/opt/auth/test/user.keytab**.
- Common mode:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

jdbc:hive2://
<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=**sparkthriftserver2x**;

For example, when you use Beeline client for connection in common mode, run the following command:

sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://***<zkNode1_IP>:<zkNode1_Port>,<zkNode2_IP>:<zkNode2_Port>,<zkNode3_IP>:<zkNode3_Port>***/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;"**

Non-multi-tenant mode

In non-multi-tenant mode, a client connects to a specified JDBCServer node. Compared with multi-active instance mode, the connection string in non-multi-active instance mode does not contain **serviceDiscoveryMode** and **zooKeeperNamespace** parameters about ZooKeeper.

For example, when you use Beeline client to connect JDBCServer in non-multi-tenant instance mode, run the following command:

sh *CLIENT_HOME*/**spark/bin/beeline -u "jdbc:hive2://***<server_IP>:<server_Port>***/;user.principal=spark/hadoop.***<System domain name>@<System domain name>***;saslQop=auth-conf;auth=KERBEROS;principal=spark/hadoop.***<System domain name>@<System domain name>***;"**

NOTE

- **<server_IP>:<server_Port>** indicates the URL of the specified JDBCServer node.
- **CLIENT_HOME** indicates the client path.

Except the connection method, other operations of JDBCServer API in multi-tenant mode and non-multi-tenant mode are the same.

Specifying a Tenant

Generally, the client submitted by a user connects to the default JDBCServer of the tenant to which the user belongs. If you want to connect the client to the JDBCServer of a specified tenant, add the **--hiveconf mapreduce.job.queuename** parameter.

Command for connecting Beeline is as follows (**aaa** indicates the tenant name):

beeline --hiveconf mapreduce.job.queuename=aaa -u 'jdbc:hive2://192.168.39.30:2181,192.168.40.210:2181,192.168.215.97:2181;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=sparkthriftserver2x;saslQop=auth-conf;auth=KERBEROS;principal=spark2x/hadoop.*<System domain name>@<System domain name>*;'

1.6.30.4 Relationship Between Spark2x and Other Components

Relationship Between Spark and HDFS

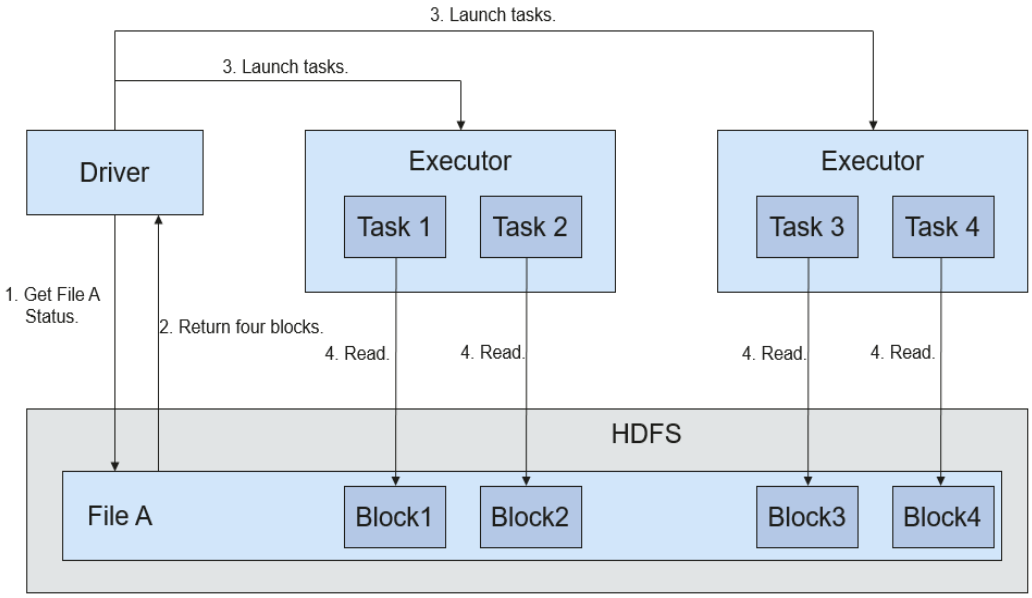
Data computed by Spark comes from multiple data sources, such as local files and HDFS. Most data comes from HDFS which can read data in large scale for parallel computing After being computed, data can be stored in HDFS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Spark involves Driver and Executor. Driver schedules tasks while Executor runs tasks.

Figure 1-126 describes the file reading process.

Figure 1-126 File reading process



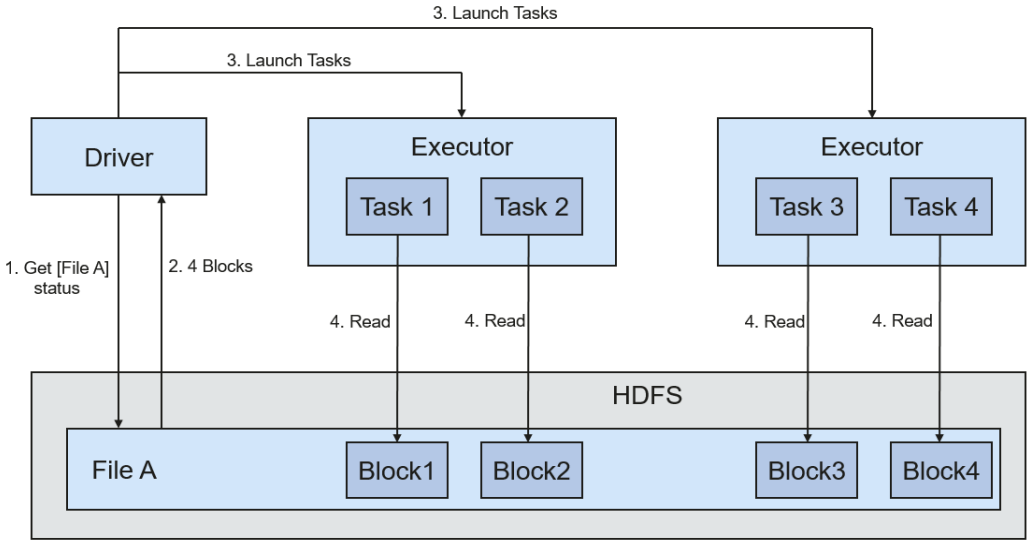
The file reading process is as follows:

1. Driver interconnects with HDFS to obtain the information of File A.
2. The HDFS returns the detailed block information about this file.
3. Driver sets a parallel degree based on the block data amount, and creates multiple tasks to read the blocks of this file.
4. Executor runs the tasks and reads the detailed blocks as part of the Resilient Distributed Dataset (RDD).

Figure 1-127 describes the file writing process.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-127 File writing process



The file writing process is as follows:

1. Driver creates a directory where the file is to be written.
2. Based on the RDD distribution status, the number of tasks related to data writing is computed, and these tasks are sent to Executor.
3. Executor runs these tasks, and writes the RDD data to the directory created in 1.

Relationship Between Spark and YARN

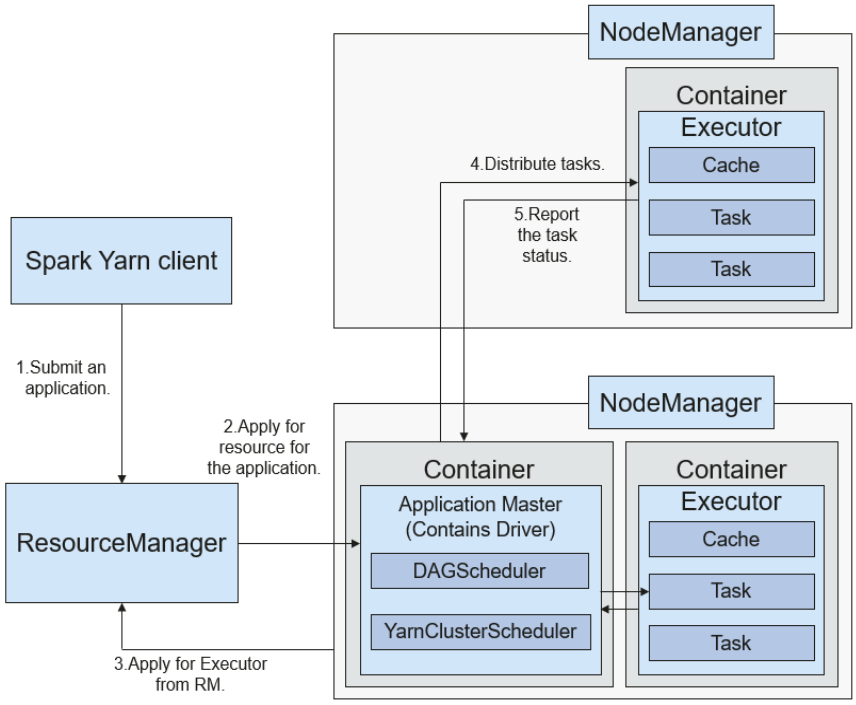
The Spark computing and scheduling can be implemented using Yarn mode. Spark enjoys the computing resources provided by Yarn clusters and runs tasks in a distributed way. Spark on Yarn has two modes: Yarn-cluster and Yarn-client.

- Yarn-cluster mode

  Figure 1-128 describes the operation framework.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-128 Spark on Yarn-cluster operation framework



Spark on Yarn-cluster implementation process:

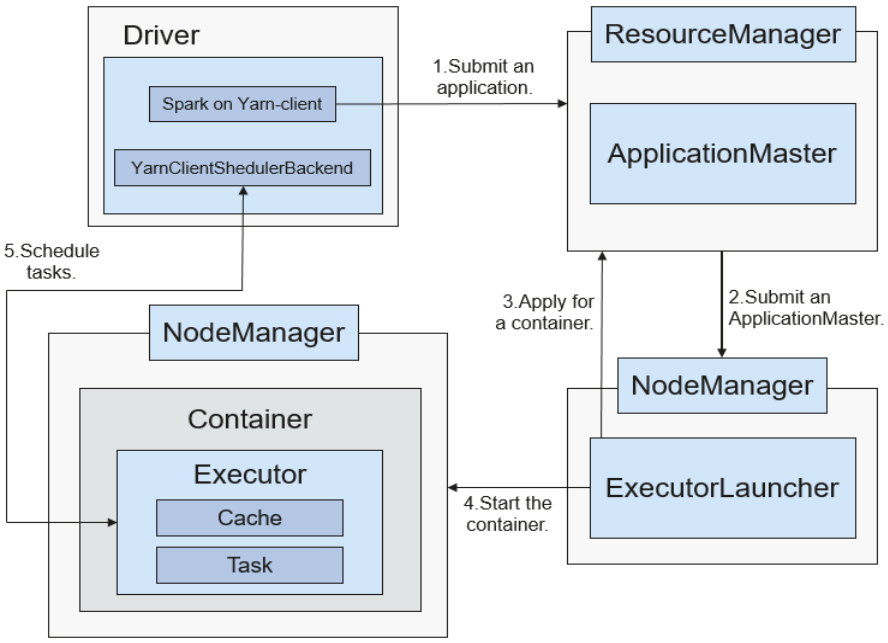
a. The client generates the application information, and then sends the information to ResourceManager.
b. ResourceManager allocates the first container (ApplicationMaster) to SparkApplication and starts the driver on the container.
c. ApplicationMaster applies for resources from ResourceManager to run the container.

   ResourceManager allocates the containers to ApplicationMaster, which communicates with the related NodeManagers and starts the executor in the obtained container. After the executor is started, it registers with drivers and applies for tasks.
d. Drivers allocate tasks to the executors.
e. Executors run tasks and report the operating status to Drivers.

- Yarn-client mode

  Figure 1-129 describes the operation framework.

Figure 1-129 Spark on Yarn-client operation framework



Spark on Yarn-client implementation process:

NOTE

In Yarn-client mode, the Driver is deployed and started on the client. In Yarn-client mode, the client of an earlier version is incompatible. The Yarn-cluster mode is recommended.

a. The client sends the Spark application request to ResourceManager, and packages all information required to start ApplicationMaster and sends the information to ResourceManager. ResourceManager then returns the results to the client. The results include information such as ApplicationId, and the upper limit as well as lower limit of available resources. After receiving the request, ResourceManager finds a proper node for ApplicationMaster and starts it on this node. ApplicationMaster is a role in Yarn, and the process name in Spark is ExecutorLauncher.

b. Based on the resource requirements of each task, ApplicationMaster can apply for a series of containers to run tasks from ResourceManager.

c. After receiving the newly allocated container list (from ResourceManager), ApplicationMaster sends information to the related NodeManagers to start the containers.

   ResourceManager allocates the containers to ApplicationMaster, which communicates with the related NodeManagers and starts the executor in the obtained container. After the executor is started, it registers with drivers and applies for tasks.

   NOTE

   Running Containers will not be suspended to release resources.

d. Drivers allocate tasks to the executors. Executors run tasks and report the operating status to Drivers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.30.5 Spark2x Open Source New Features

Compared with Spark 1.5, Spark2*x* has some new open-source features.

The specific features or concepts are as follows:

- DataSet: For details, see **SparkSQL and DataSet Principle**.
- Spark SQL Native DDL/DML: For details, see **SparkSQL and DataSet Principle**.
- SparkSession: For details, see **SparkSession Principle**.
- Structured Streaming: For details, see **Structured Streaming Principle**.
- Optimizing Small Files
- Optimizing the Aggregate Algorithm
- Optimizing Datasource Tables
- Merging CBO

1.6.30.6 Optimizing Spark SQL Query of Data of Multiple Sources

Scenario

Enterprises usually store massive data, such as from various databases and warehouses, for management and information collection. However, diversified data sources, hybrid dataset structures, and scattered data storage lower query efficiency.

The open source Spark only supports simple filter pushdown during querying of multi-source data. The SQL engine performance is deteriorated due of a large amount of unnecessary data transmission. The pushdown function is enhanced, so that **aggregate**, complex **projection**, and complex **predicate** can be pushed to data sources, reducing unnecessary data transmission and improving query performance.

Only the JDBC data source supports pushdown of query operations, such as **aggregate**, **projection**, **predicate**, **aggregate over inner join**, and **aggregate over union all**. All pushdown operations can be enabled based on your requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-27 Enhanced query of cross-source query

Module	Before Enhancement	After Enhancement
aggregate	The pushdown of **aggregate** is not supported.	• Aggregation functions including **sum**, **avg**, **max**, **min**, and **count** are supported. Example: select count(*) from table • Internal expressions of aggregation functions are supported. Example: select sum(a+b) from table • Calculation of aggregation functions is supported. Example: select avg(a) + max(b) from table • Pushdown of **having** is supported. Example: select sum(a) from table where a>0 group by b having sum(a)>10 • Pushdown of some functions is supported. Pushdown of lines in mathematics, time, and string functions, such as **abs()**, **month()**, and **length()** are supported. In addition to the preceding built-in functions, you can run the **SET** command to add functions supported by data sources. Example: select sum(abs(a)) from table • Pushdown of **limit** and **order by** after **aggregate** is supported. However, the pushdown is not supported in Oracle, because Oracle does not support **limit**. Example: select sum(a) from table where a>0 group by b order by sum(a) limit 5
projection	Only pushdown of simple **projection** is supported. Example: select a, b from table	• Complex expressions can be pushed down. Example: select (a+b)*c from table • Some functions can be pushed down. For details, see the description below the table. Example: select length(a)+abs(b) from table • Pushdown of **limit** and **order by** after **projection** is supported. Example: select a, b+c from table order by a limit 3

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Module	Before Enhancement	After Enhancement
predicate	Only simple filtering with the column name on the left of the operator and values on the right is supported. Example: select * from table where a>0 or b in ("aaa", "bbb")	• Complex expression pushdown is supported. Example: select * from table where a +b>c*d or a/c in (1, 2, 3) • Some functions can be pushed down. For details, see the description below the table. Example: select * from table where length(a)>5
aggregate over inner join	Related data from the two tables must be loaded to Spark. The join operation must be performed before the **aggregate** operation.	The following functions are supported: • Aggregation functions including **sum**, **avg**, **max**, **min**, and **count** are supported. • All **aggregate** operations can be performed in a same table. The **group by** operations can be performed on one or two tables and only inner join is supported. The following scenarios are not supported: • **aggregate** cannot be pushed down from both the left- and right-join tables. • **aggregate** contains operations, for example, sum(a+b). • **aggregate** operations, for example, sum(a)+min(b).
aggregate over union all	Related data from the two tables must be loaded to Spark. **union** must be performed before **aggregate**.	Supported scenarios: Aggregation functions including **sum**, **avg**, **max**, **min**, and **count** are supported. Unsupported scenarios: • **aggregate** contains operations, for example, sum(a+b). • **aggregate** operations, for example, sum(a)+min(b).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Precautions

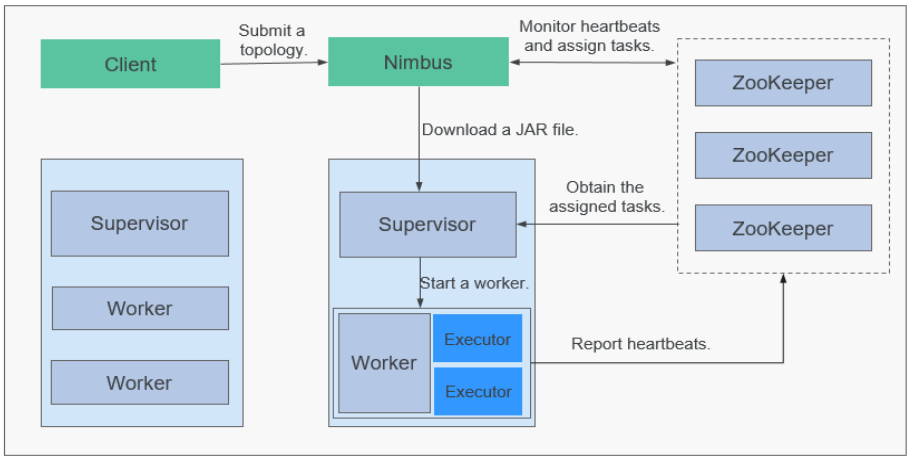
- If external data source is Hive, query operation cannot be performed on foreign tables created by Spark.
- Only MySQL and MPPDB data sources are supported.

1.6.31 Storm

1.6.31.1 Storm Basic Principles

Apache Storm is a distributed, reliable, and fault-tolerant real-time stream data processing system. In Storm, a graph-shaped data structure called topology needs to be designed first for real-time computing. The topology will be submitted to a cluster. Then a master node in the cluster distributes codes and assigns tasks to worker nodes. A topology contains two roles: spout and bolt. A spout sends messages and sends data streams in tuples. A bolt converts the data streams and performs computing and filtering operations. The bolt can randomly send data to other bolts. Tuples sent by a spout are unchangeable arrays and map to fixed key-value pairs.

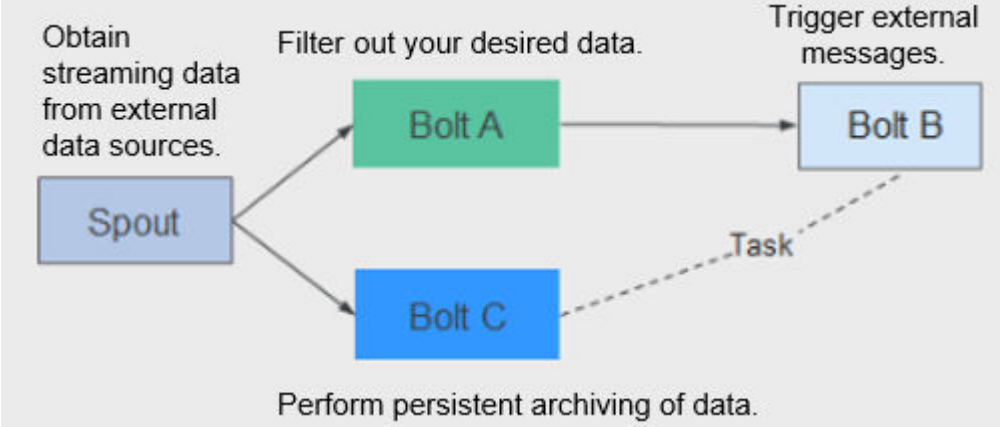
Figure 1-130 Basic architecture of Storm



Service processing logic is encapsulated in the topology of Storm. A topology is a set of spout (data sources) and bolt (logical processing) components that are connected using Stream Groupings in DAG mode. All components (spout and bolt) in a topology are working in parallel. In a topology, you can specify the parallelism for each node. Then, Storm allocates tasks in the cluster for computing to improve system processing capabilities.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-131 Topology



Storm is applicable to real-time analysis, continuous computing, and distributed extract, transform, and load (ETL). It has the following advantages:

- Wide applications
- High scalability
- Zero data loss
- High fault tolerance
- Easy to construct and control
- Multi-language support

Storm is a computing platform and provides Continuous Query Language (CQL) in the service layer to facilitate service implementation. CQL has the following features:

- Easy to use: The CQL syntax is similar to the SQL syntax. Users who have basic knowledge of SQL can easily learn CQL and use it to develop services.
- Rich functions: In addition to basic expressions provided by SQL, CQL provides functions, such as windows, filtering, and concurrency setting, for stream processing.
- Easy to scale: CQL provides an extension API to support increasingly complex service scenarios. Users can customize the input, output, serialization, and deserialization to meet specific service requirements.
- Easy to debug: CQL provides detailed explanation of error codes, facilitating users to rectify faults.

For details about Storm architecture and principles, see **https://storm.apache.org/**.

Principle

- **Basic Concepts**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

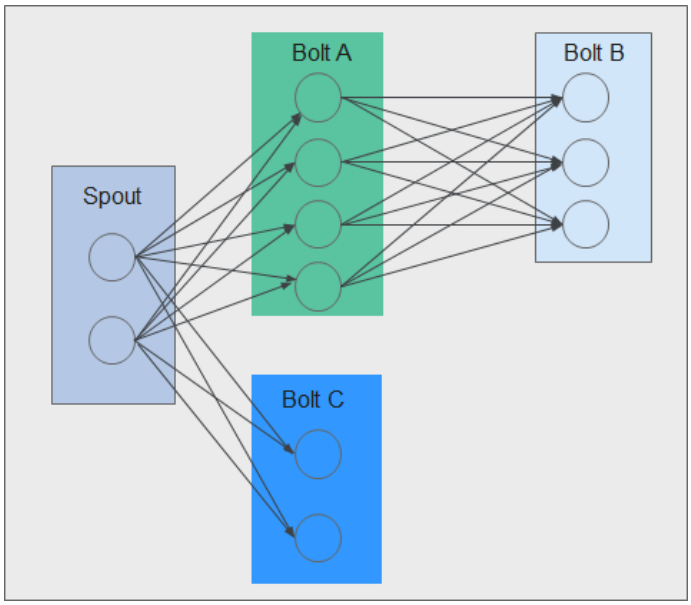
Table 1-28 Concepts

Concept	Description
Tuple	A tuple is an invariable key-value pair used to transfer data. Tuples are created and processed in distributed manner.
Stream	A stream is an unbounded sequence of tuples.
Topology	A topology is a real-time application running on the Storm platform. It is a Directed Acyclic Graph (DAG) composed of components. A topology can concurrently run on multiple machines. Each machine runs a part of the DAG. A topology is similar to a MapReduce job. The difference is that the topology is a resident program. Once started, the topology cannot stop unless it is manually terminated.
Spout	A spout is the source of tuples. For example, a spout may read data from a message queue, database, file system, or TCP connection and converts them as tuples, which are processed by the next component.
Bolt	In a Topology, a bolt is a component that receives data and executes specific logic, such as filtering or converting tuples, joining or aggregating streams, and performing statistics and result persistence.
Worker	A Worker is a physical processing in running state in a Topology. Each Worker is a JVM process. Each Topology may be executed by multiple Workers. Each Worker executes a logic subset of the Topology.
Task	A task is a spout or bolt thread of a Worker.
Stream groupings	A stream grouping specifies the tuple dispatching policies. It instructs the subsequent bolt how to receive tuples. The supported policies include Shuffle Grouping, Fields Grouping, All Grouping, Global Grouping, Non Grouping, and Directed Grouping.

Figure 1-132 shows a Topology (DAG) consisting of a Spout and Bolt. In the figure, a rectangle indicates a Spout or Bolt, the nodes in each rectangle indicate tasks, and the lines between tasks indicate streams.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-132 Topology



- **Reliability**

  Storm provides three levels of data reliability:

  - At Most Once: The processed data may be lost, but it cannot be processed repeatedly. This reliability level offers the highest throughput.
  - At Least Once: Data may be processed repeatedly to ensure reliable data transmission. If a response is not received within the specified time, the Spout resends the data to Bolts for processing. This reliability level may slightly affect system performance.
  - Exactly Once: Data is successfully transmitted without loss or redundancy processing. This reliability level delivers the poorest performance.

  Select the reliability level based on service requirements. For example, for the services requiring high data reliability, use Exactly Once to ensure that data is processed only once. For the services insensitive to data loss, use other levels to improve system performance.

- **Fault Tolerance**

  Storm is a fault-tolerant system that offers high availability. **Table 1-29** describes the fault tolerance of the Storm components.

Table 1-29 Fault tolerance

Scenario	Description
Nimbus failed	Nimbus is fail-fast and stateless. If the active Nimbus is faulty, the standby Nimbus takes over services immediately, and provide external services.
Supervisor failed	Supervisor is a background daemon of Workers. It is fail-fast and stateless. If a Supervisor is faulty, the Workers running on the node are not affected but cannot receive new tasks. The OMS can detect the fault of the Supervisor and restart the processes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scenario	Description
Worker failed	If a Worker is faulty, the Supervisor on the Worker will restart it again. If the restart fails for multiple times, Nimbus reassigns tasks to other nodes.
Node failed	If a node is faulty, all the tasks being processed by the node time out and Nimbus will assign the tasks to another node for processing.

Open Source Features

- Distributed real-time computing

  In a Storm cluster, each machine supports the running of multiple work processes and each work process can create multiple threads. Each thread can execute multiple tasks. A task indicates concurrent data processing.

- High fault tolerance

  During message processing, if a node or a process is faulty, the message processing unit can be redeployed.

- Reliable messages

  Data processing methods including At-Least Once, At-Most Once, and Exactly Once are supported.

- Security mechanism

  Storm provides Kerberos-based authentication and pluggable authorization mechanisms, supports SSL Storm UI and Log Viewer UI, and supports security integration with other big data platform components (such as ZooKeeper and HDFS).

- Flexible topology defining and deployment

  The Flux framework is used to define and deploy service topologies. If the service DAG is changed, users only need to modify YAML domain specific language (DSL), but do not need to recompile or package service code.

- Integration with external components

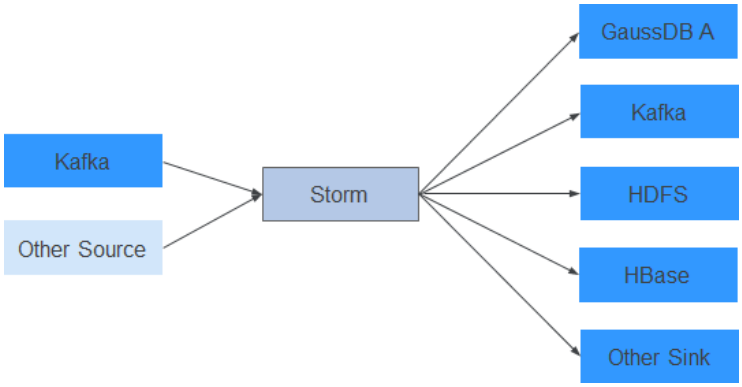
  Storm supports integration with multiple external components such as Kafka, HDFS, HBase, Redis, and JDBC/RDBMS, implementing services that involve multiple data sources.

1.6.31.2 Relationship Between Storm and Other Components

Storm provides a real-time distributed computing framework. It can obtain real-time messages from data sources (such as Kafka and TCP connection), perform high-throughput and low-latency real-time computing on a real-time platform, and export results to message queues or implement data persistence. **Figure 1-133** shows the relationship between Storm and other components.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-133 Relationship with other components



Relationship between Storm and Streaming

Both Storm and Streaming use the open source Apache Storm kernel. However, the kernel version used by Storm is 1.2.1 whereas that used by Streaming is 0.10.0. Streaming is used to inherit transition services in upgrade scenarios. For example, if Streaming has been deployed and services are running, Streaming can still be used after an upgrade. Storm is recommended in a new cluster.

Storm 1.2.1 has the following new features:

- **Distributed cache**: Provides external resources (configurations) required for sharing and updating the topology using CLI tools. You do not need to re-package and re-deploy the topology.
- **Native Streaming Window API**: Provides window-based APIs.
- **Resource scheduler**: Added the resource scheduler plug-in. When defining a topology, you can specify the maximum resources available and assign resource quotas to users, thus to manage topology resources of the users.
- **State management**: Provides the Bolt API with the checkpoint mechanism. When an event fails, Storm automatically manages the Bolt status and restore the event.
- **Message sampling and debugging**: On the Storm UI, you can enable or disable topology- or component-level debugging to output stream messages to specified logs based on the sampling ratio.
- **Worker dynamic analysis**: On the Storm UI, you can collect jstack and heap logs of the Worker process and restart the Worker process.
- **Dynamic adjustment of topology logs**: You can dynamically change the running topology logs on the CLI or Storm UI.
- **Improved performance**: Compared with earlier versions, the performance of Storm is greatly improved. Although the topology performance is closely related to the use case scenario and dependency on external services, the performance is three times higher in most scenarios.

1.6.31.3 Storm Enhanced Open Source Features

- CQL

  Continuous Query Language (CQL) is an SQL-like language used for real-time stream processing. Compared with SQL, CQL has introduced the concept of (time-sequencing) window, which allows data to be stored and processed in

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the memory. The CQL output is the computing results of data streams at specific time. The use of CQL accelerates service development, enables tasks to be easily submitted to the Storm platform for real-time processing, facilitates output of results, and allows tasks to be terminated at the appropriate time.

- High Availability

  Nimbus HA ensures continuous service processing such as adding topologies and management even if one Nimbus is faulty, improving cluster availability.

1.6.32 Tez

Tez is Apache's latest open source computing framework that supports Directed Acyclic Graph (DAG) jobs. It can convert multiple dependent jobs into one job, greatly improving the performance of DAG jobs.

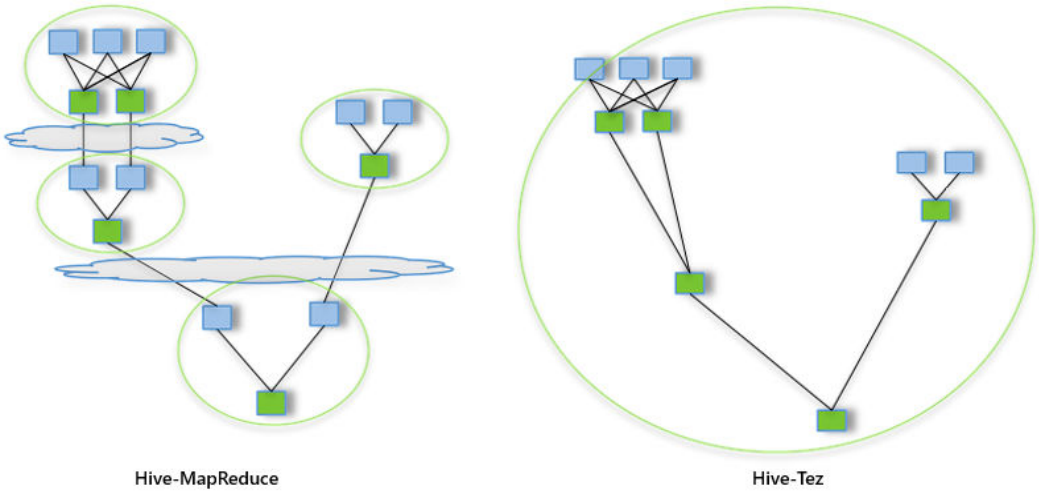
MRS uses Tez as the default execution engine of Hive. Tez remarkably surpasses the original MapReduce computing engine in terms of execution efficiency.

For details about Tez, see **https://tez.apache.org/**.

Relationship Between Tez and MapReduce

Tez uses a DAG to organize MapReduce tasks. In the DAG, a node is an RDD, and an edge indicates an operation on the RDD. The core idea is to further split Map tasks and Reduce tasks. A Map task is split into the Input-Processor-Sort-Merge-Output tasks, and the Reduce task is split into the Input-Shuffle-Sort-Merge-Process-Output tasks. Tez flexibly regroups several small tasks to form a large DAG job.

Figure 1-134 Processes for submitting tasks using Hive on MapReduce and Hive on Tez



A Hive on MapReduce task contains multiple MapReduce tasks. Each task stores intermediate results to HDFS. The reducer in the previous step provides data for the mapper in the next step. A Hive on Tez task can complete the same processing process in only one task, and HDFS does not need to be accessed between tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Relationship Between Tez and Yarn

Tez is a computing framework running on Yarn. The runtime environment consists of ResourceManager and ApplicationMaster of Yarn. ResourceManager is a brand new resource manager system, and ApplicationMaster is responsible for cutting MapReduce job data, assigning tasks, applying for resources, scheduling tasks, and tolerating faults. In addition, TezUI depends on TimelineServer provided by Yarn to display the running process of Tez tasks.

1.6.33 YARN

1.6.33.1 YARN Basic Principles

The Apache open source community introduces the unified resource management framework YARN to share Hadoop clusters, improve their scalability and reliability, and eliminate a performance bottleneck of JobTracker in the early MapReduce framework.

The fundamental idea of YARN is to split up the two major functionalities of the JobTracker, resource management and job scheduling/monitoring, into separate daemons. The idea is to have a global ResourceManager (RM) and per-application ApplicationMaster (AM).

NOTE

- To use Yarn, ensure that the Hadoop service has been installed in the MRS cluster.
- An application is either a single job in the classical sense of MapReduce jobs or a Directed Acyclic Graph (DAG) of jobs.

Architecture

ResourceManager is the essence of the layered structure of YARN. This entity controls an entire cluster and manages the allocation of applications to underlying compute resources. The ResourceManager carefully allocates various resources (compute, memory, bandwidth, and so on) to underlying NodeManagers (YARN's per-node agents). The ResourceManager also works with ApplicationMasters to allocate resources, and works with the NodeManagers to start and monitor their underlying applications. In this context, the ApplicationMaster has taken some of the role of the prior TaskTracker, and the ResourceManager has taken the role of the JobTracker.

ApplicationMaster manages each instance of an application running in YARN. The ApplicationMaster negotiates resources from the ResourceManager and works with the NodeManagers to monitor container execution and resource usage (CPU and memory resource allocation).

The NodeManager manages each node in a YARN cluster. The NodeManager provides per-node services in a cluster, from overseeing the management of a container over its lifecycle to monitoring resources and tracking the health of its nodes. MRv1 manages execution of the Map and Reduce tasks through slots, whereas the NodeManager manages abstract containers, which represent per-node resources available for a particular application.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-135 Architecture

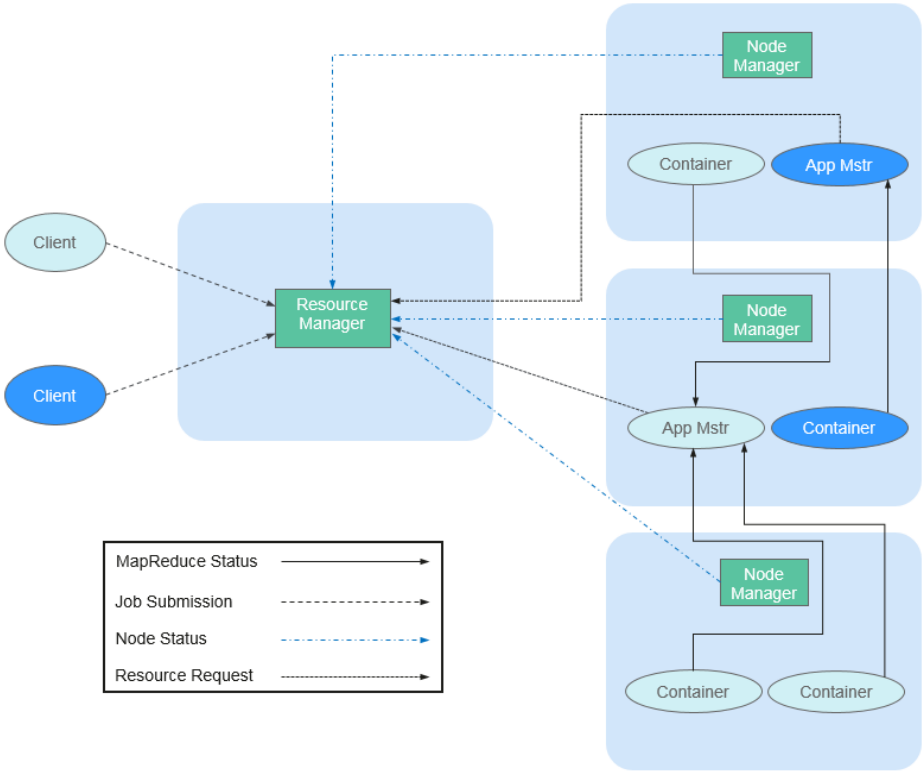


Table 1-30 describes the components shown in Figure 1-135.

Table 1-30 Architecture description

Name	Description
Client	Client of a YARN application. You can submit a task to ResourceManager and query the operating status of an application using the client.
ResourceManager(RM)	RM centrally manages and allocates all resources in the cluster. It receives resource reporting information from each node (NodeManager) and allocates resources to applications on the basis of the collected resources according a specified policy.
NodeManager(NM)	The NodeManager (NM) is the agent running on each YARN node. It manages resources of a compute node in the Hadoop cluster, communicates with the RM, tracks container lifecycles, monitors resource usage like memory and CPU for each container, checks node health, and handles logs and auxiliary services of applications.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

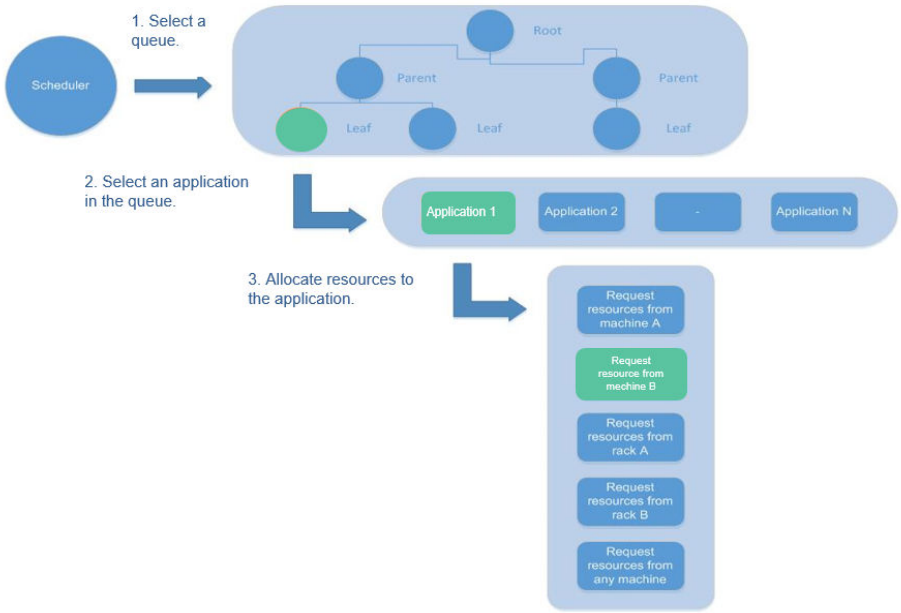
Name	Description
ApplicationMaster(AM)	AM (App Mstr in the figure above) is responsible for all tasks through the lifecycle of in an application. The tasks include the following: Negotiate with an RM scheduler to obtain a resource; further allocate the obtained resources to internal tasks (secondary allocation of resources); communicate with the NM to start or stop tasks; monitor the running status of all tasks; and apply for resources for tasks again to restart the tasks when the tasks fail to be executed.
Container	A resource abstraction in YARN. It encapsulates multi-dimensional resources (including only memory and CPU) on a certain node. When ApplicationMaster applies for resources from ResourceManager, the ResourceManager returns resources to the ApplicationMaster in a container. YARN allocates one container for each task and the task can only use the resources encapsulated in the container.

In YARN, resource schedulers organize resources through hierarchical queues. This ensures that resources are allocated and shared among queues, thereby improving the usage of cluster resources. The core resource allocation model of Superior Scheduler is the same as that of Capacity Scheduler, as shown in the following figure.

A scheduler maintains queue information. You can submit applications to one or more queues. During each NM heartbeat, the scheduler selects a queue according to a specific scheduling rule, selects an application in the queue, and then allocates resources to the application. If resources fail to be allocated to the application due to the limit of some parameters, the scheduler will select another application. After the selection, the scheduler processes the resource request of this application. The scheduler gives priority to the requests for local resources first, and then for resources on the same rack, and finally for resources from any machine.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-136 Resource allocation model



Principle

The new Hadoop MapReduce framework is named MRv2 or YARN. YARN consists of ResourceManager, ApplicationMaster, and NodeManager.

- ResourceManager is a global resource manager that manages and allocates resources in the system. ResourceManager consists of Scheduler and Applications Manager.
  - Scheduler allocates system resources to all running applications based on the restrictions such as capacity and queue (for example, allocates a certain amount of resources for a queue and executes a specific number of jobs). It allocates resources based on the demand of applications, with container being used as the resource allocation unit. Functioning as a dynamic resource allocation unit, Container encapsulates memory, CPU, disk, and network resources, thereby limiting the resource consumed by each task. In addition, the Scheduler is a pluggable component. You can design new schedulers as required. YARN provides multiple directly available schedulers, such as Fair Scheduler and Capacity Scheduler.
  - Applications Manager manages all applications in the system and involves submitting applications, negotiating with schedulers about resources, enabling and monitoring ApplicationMaster, and restarting ApplicationMaster upon the startup failure.
- NodeManager is the resource and task manager of each node. On one hand, NodeManager periodically reports resource usage of the local node and the running status of each Container to ResourceManager. On the other hand, NodeManager receives and processes requests from ApplicationMaster for starting or stopping Containers.
- ApplicationMaster is responsible for all tasks through the lifecycle of an application, these channels include the following:
  - Negotiate with the RM scheduler to obtain resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Assign resources to internal components (secondary allocation of resources).
  - Communicates with NodeManager to start or stop tasks.
  - Monitor the running status of all tasks, and applies for resources again for tasks when tasks fail to run to restart the tasks.

Capacity Scheduler Principle

Capacity Scheduler is a multi-user scheduler. It allocates resources by queue and sets the minimum/maximum resources that can be used for each queue. In addition, the upper limit of resource usage is set for each user to prevent resource abuse. Remaining resources of a queue can be temporarily shared with other queues.

Capacity Scheduler supports multiple queues. It configures a certain amount of resources for each queue and adopts the first-in-first-out queuing (FIFO) scheduling policy. To prevent one user's applications from exclusively using the resources in a queue, Capacity Scheduler sets a limit on the number of resources used by jobs submitted by one user. During scheduling, Capacity Scheduler first calculates the number of resources required for each queue, and selects the queue that requires the least resources. Then, it allocates resources based on the job priority and time that jobs are submitted as well as the limit on resources and memory. Capacity Scheduler supports the following features:

- Guaranteed capacity: As the MRS cluster administrator, you can set the lower and upper limits of resource usage for each queue. All applications submitted to this queue share the resources.
- High flexibility: Temporarily, the remaining resources of a queue can be shared with other queues. However, such resources must be released in case of new application submission to the queue. Such flexible resource allocation helps notably improve resource usage.
- Multi-tenancy: Multiple users can share a cluster, and multiple applications can run concurrently. To avoid exclusive resource usage by a single application, user, or queue, the MRS cluster administrator can add multiple constraints (for example, limit on concurrent tasks of a single application).
- Assured protection: An ACL list is provided for each queue to strictly limit user access. You can specify the users who can view your application status or control the applications. Additionally, the MRS cluster administrator can specify a queue administrator and a cluster system administrator.
- Dynamic update of configuration files: MRS cluster administrators can dynamically modify configuration parameters to manage clusters online.

Each queue in Capacity Scheduler can limit the resource usage. However, the resource usage of a queue determines its priority when resources are allocated to queues, indicating that queues with smaller capacity are competitive. If the throughput of a cluster is big, delay scheduling enables an application to give up cross-machine or cross-rack scheduling, and to request local scheduling.

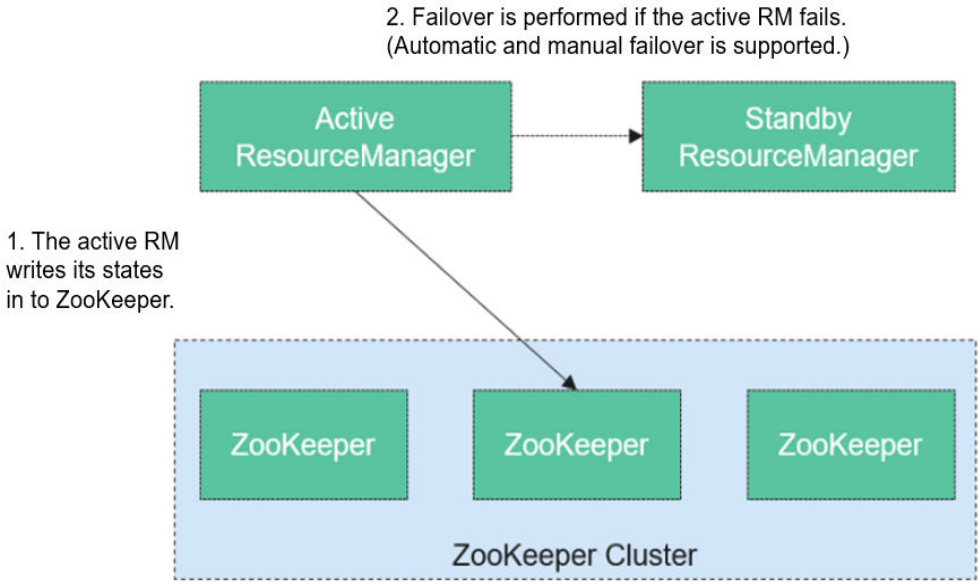
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.6.33.2 YARN HA Solution

HA Principles and Implementation Solution

ResourceManager in YARN manages resources and schedules tasks in the cluster. In versions earlier than Hadoop 2.4, SPOFs may occur on ResourceManager in the YARN cluster. The YARN HA solution uses redundant ResourceManager nodes to tackle challenges of service reliability and fault tolerance.

Figure 1-137 ResourceManager HA architecture



ResourceManager HA is achieved using active-standby ResourceManager nodes, as shown in **Figure 1-137**. Similar to the HDFS HA solution, the ResourceManager HA allows only one ResourceManager node to be in the active state at any time. When the active ResourceManager fails, the active-standby switchover can be triggered automatically or manually.

When the automatic failover function is not enabled, after the YARN cluster is enabled, MRS cluster administrators need to run the ***yarn rmadmin*** command to manually switch one of the ResourceManager nodes to the active state. Upon a planned maintenance event or a fault, they are expected to first demote the active ResourceManager to the standby state and the standby ResourceManager promote to the active state.

When automatic failover is enabled, a built-in ActiveStandbyElector that is based on ZooKeeper is used to decide which ResourceManager node should be the active one. When the active ResourceManager is faulty, another ResourceManager node is automatically selected to be the active one to take over the faulty node.

When ResourceManager nodes in the cluster are deployed in HA mode, the configuration **yarn-site.xml** used by clients needs to list all the ResourceManager nodes. The client (including ApplicationMaster and NodeManager) searches for the active ResourceManager in polling mode. That is, the client needs to provide the fault tolerance mechanism. If the active ResourceManager cannot be connected with, the client continuously searches for a new one in polling mode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the standby ResourceManager node becomes active, the upper-layer applications can recover to their status when the fault occurred. When ResourceManager Restart is enabled, the restarted ResourceManager node loads the information of the previous active ResourceManager node, and takes over container status information on all NodeManager nodes to continue service running. In this way, status information can be saved by periodically executing checkpoint operations, avoiding data loss. Ensure that both active and standby ResourceManager nodes can access the status information. Currently, three methods are provided for sharing status information by file system (FileSystemRMStateStore), LevelDB database (LeveldbRMStateStore), and ZooKeeper (ZKRMStateStore). Among them, only ZKRMStateStore supports the Fencing mechanism. By default, Hadoop uses ZKRMStateStore.

1.6.33.3 Relationship Between YARN and Other Components

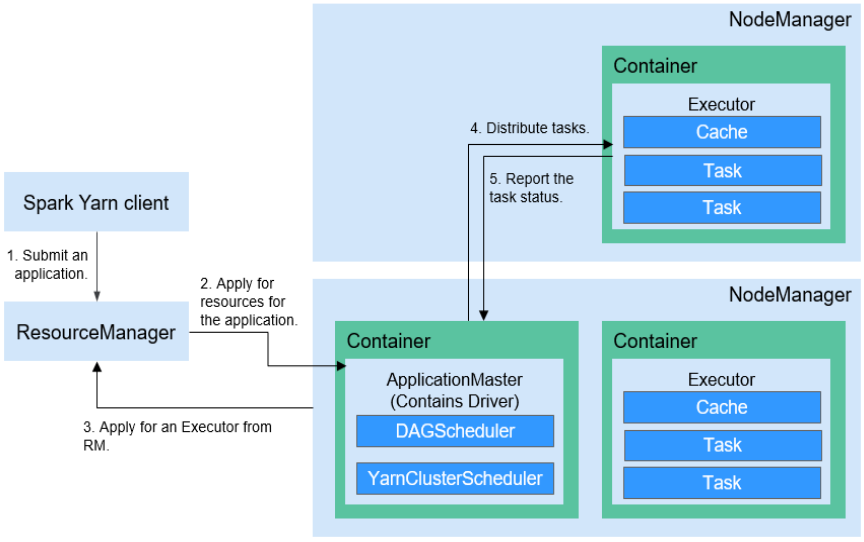
Relationship Between YARN and Spark

The Spark computing and scheduling can be implemented using YARN mode. Spark enjoys the compute resources provided by YARN clusters and runs tasks in a distributed way. Spark on YARN has two modes: YARN-cluster and YARN-client.

- YARN Cluster mode

  Figure 1-138 describes the operation framework.

  Figure 1-138 Spark on YARN-cluster operation framework



  Spark on YARN-cluster implementation process:

  a. The client generates the application information, and then sends the information to ResourceManager.
  b. ResourceManager allocates the first container (ApplicationMaster) to SparkApplication and starts the driver on the container.
  c. ApplicationMaster applies for resources from ResourceManager to run the container.

     ResourceManager allocates the containers to ApplicationMaster, which communicates with the related NodeManagers and starts the executor in

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the obtained container. After the executor is started, it registers with drivers and applies for tasks.

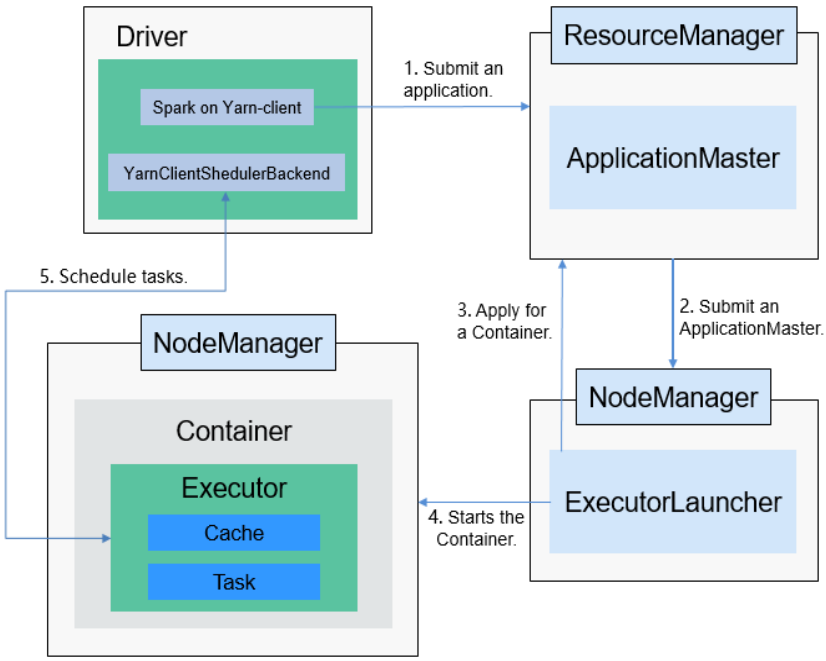
d. Drivers allocate tasks to the executors.

e. Executors run tasks and report the operating status to Drivers.

- YARN Client mode

  Figure 1-139 describes the operation framework.

  Figure 1-139 Spark on YARN-client operation framework



Spark on YARN-client implementation process:

NOTE

In YARN-client mode, the driver is deployed and started on the client. In YARN-client mode, the client of an earlier version is incompatible. You are advised to use the YARN-cluster mode.

a. The client sends the Spark application request to ResourceManager, then ResourceManager returns the results. The results include information such as Application ID and the maximum and minimum available resources. The client packages all information required to start ApplicationMaster, and sends the information to ResourceManager.

b. After receiving the request, ResourceManager finds a proper node for ApplicationMaster and starts it on this node. ApplicationMaster is a role in YARN, and the process name in Spark is ExecutorLauncher.

c. Based on the resource requirements of each task, ApplicationMaster can apply for a series of containers to run tasks from ResourceManager.

d. After receiving the newly allocated container list (from ResourceManager), ApplicationMaster sends information to the related NodeManagers to start the containers.

   ResourceManager allocates the containers to ApplicationMaster, which communicates with the related NodeManagers and starts the executor in

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the obtained container. After the executor is started, it registers with drivers and applies for tasks.

NOTE

Running containers are not suspended and resources are not released.

e. Drivers allocate tasks to the executors. Executors run tasks and report the operating status to Drivers.

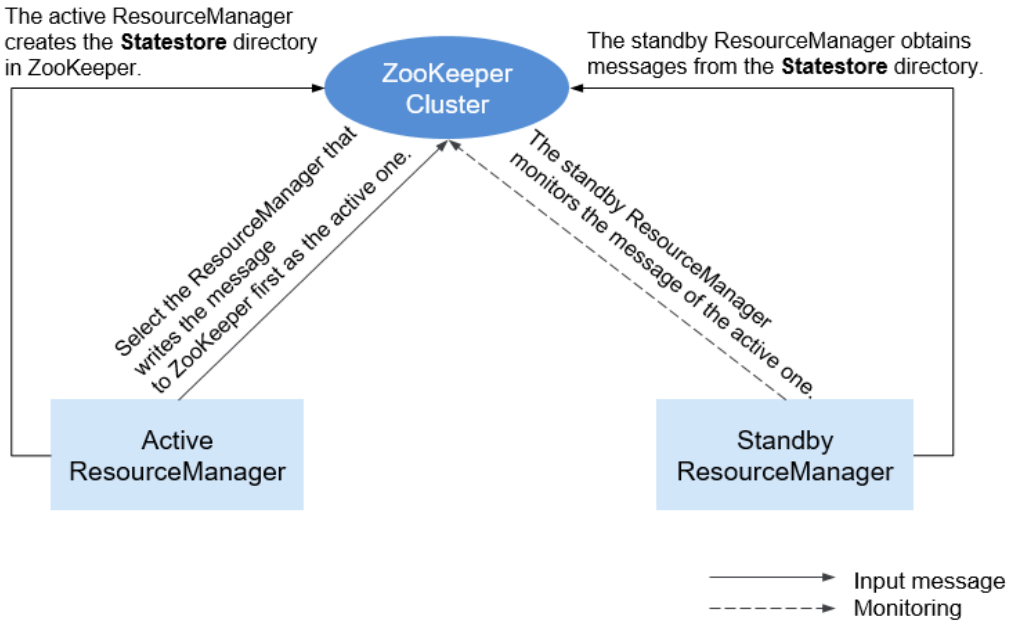
Relationship Between YARN and MapReduce

MapReduce is a computing framework running on YARN, which is used for batch processing. MRv1 is implemented based on MapReduce in Hadoop 1.0, which is composed of programming models (new and old programming APIs), running environment (JobTracker and TaskTracker), and data processing engine (MapTask and ReduceTask). This framework is still weak in scalability, fault tolerance (JobTracker SPOF), and compatibility with multiple frameworks. (Currently, only the MapReduce computing framework is supported.) MRv2 is implemented based on MapReduce in Hadoop 2.0. The source code reuses MRv1 programming models and data processing engine implementation, and the running environment is composed of ResourceManager and ApplicationMaster. ResourceManager is a brand new resource manager system, and ApplicationMaster is responsible for cutting MapReduce job data, assigning tasks, applying for resources, scheduling tasks, and tolerating faults.

Relationship Between YARN and ZooKeeper

Figure 1-140 shows the relationship between ZooKeeper and YARN.

Figure 1-140 Relationship between ZooKeeper and YARN



1. When the system is started, ResourceManager attempts to write state information to ZooKeeper. ResourceManager that first writes state information to ZooKeeper is selected as the active ResourceManager, and

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

others are standby ResourceManagers. The standby ResourceManagers periodically monitor active ResourceManager election information in ZooKeeper.

2. The active ResourceManager creates the **Statestore** directory in ZooKeeper to store application information. If the active ResourceManager is faulty, the standby ResourceManager obtains application information from the **Statestore** directory and restores the data.

Relationship Between YARN and Tez

The Hive on Tez job information requires the TimeLine Server capability of YARN so that Hive tasks can display the current and historical status of applications, facilitating storage and retrieval.

NOTE

The TimelineServer saves data to the in-memory database LevelDB, which is memory-intensive. It is imperative to allocate a minimum of 30 GB of memory to the node hosting the TimelineServer.

1.6.33.4 Yarn Enhanced Open Source Features

Priority-based task scheduling

In the native Yarn resource scheduling mechanism, if the whole Hadoop cluster resources are occupied by those MapReduce jobs submitted earlier, jobs submitted later will be kept in pending state until all running jobs are executed and resources are released.

The MRS cluster provides the task priority scheduling mechanism. With this feature, you can define jobs of different priorities. Jobs of high priority can preempt resources released from jobs of low priority though the high-priority jobs are submitted later. The low-priority jobs that are not started will be suspended unless those jobs of high priority are completed and resources are released, then they can properly be started.

This feature enables services to control computing jobs more flexibly, thereby achieving higher cluster resource utilization.

NOTE

Container reuse may conflict with task priority scheduling. If container reuse is enabled, resources are being occupied, and task priority scheduling does not take effect.

Yarn Permission Control

The permission mechanism of Hadoop Yarn is implemented through ACLs. The following describes how to grant different permission control to different users:

- Admin ACL

  An O&M administrator is specified for the YARN cluster. The Admin ACL is determined by **yarn.admin.acl**. The cluster O&M administrator can access the ResourceManager web UI and operate NodeManager nodes, queues, and NodeLabel, **but cannot submit tasks**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Queue ACL

  To facilitate user management in the cluster, users or user groups are divided into several queues to which each user and user group belongs. Each queue contains permissions to submit and manage applications (for example, terminate any application).

Open source functions:

Currently, Yarn supports the following roles for users:

- Cluster O&M administrator
- Queue administrator
- Common user

However, the APIs (such as the web UI, REST API, and Java API) provided by Yarn do not support role-specific permission control. Therefore, all users have the permission to access the application and cluster information, which does not meet the isolation requirements in the multi-tenant scenario.

This is an enhanced function.

In security mode, permission management is enhanced for the APIs such as web UI, REST API, and Java API provided by Yarn. Permission control can be performed based on user roles.

Role-based permissions are as follows:

- Cluster O&M administrator: performs management operations in the Yarn cluster, such as accessing the ResourceManager web UI, refreshing queues, setting NodeLabel, and performing active/standby switchover.
- Queue administrator: has the permission to modify and view queues managed by the Yarn cluster.
- Common user: has the permission to modify and view self-submitted applications in the Yarn cluster.

Superior Scheduler Principle (Self-developed)

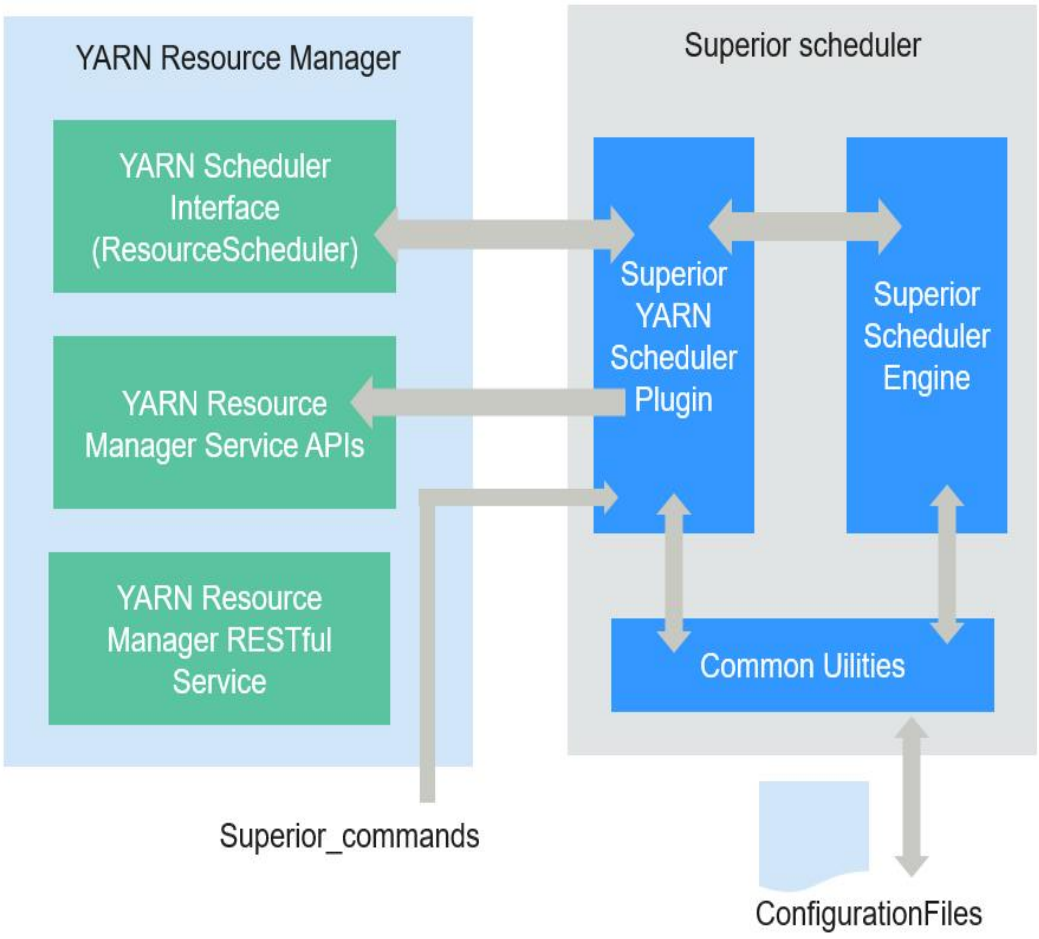
Superior Scheduler is a scheduling engine designed for the Hadoop Yarn distributed resource management system. It is a high-performance and enterprise-level scheduler designed for converged resource pools and multi-tenant service requirements.

Superior Scheduler achieves all functions of open source schedulers, Fair Scheduler, and Capacity Scheduler. Compared with the open source schedulers, Superior Scheduler is enhanced in the enterprise multi-tenant resource scheduling policy, resource isolation and sharing among users in a tenant, scheduling performance, system resource usage, and cluster scalability. Superior Scheduler is designed to replace open source schedulers.

Similar to open source Fair Scheduler and Capacity Scheduler, Superior Scheduler follows the Yarn scheduler plugin API to interact with Yarn ResourceManager to offer resource scheduling functionalities. **Figure 1-141** shows the overall system diagram.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-141 Internal architecture of Superior Scheduler



In **Figure 1-141**, Superior Scheduler consists of the following modules:

- Superior Scheduler Engine is a high performance scheduler engine with rich scheduling policies.
- Superior Yarn Scheduler Plugin functions as a bridge between Yarn ResourceManager and Superior Scheduler Engine and interacts with Yarn ResourceManager.

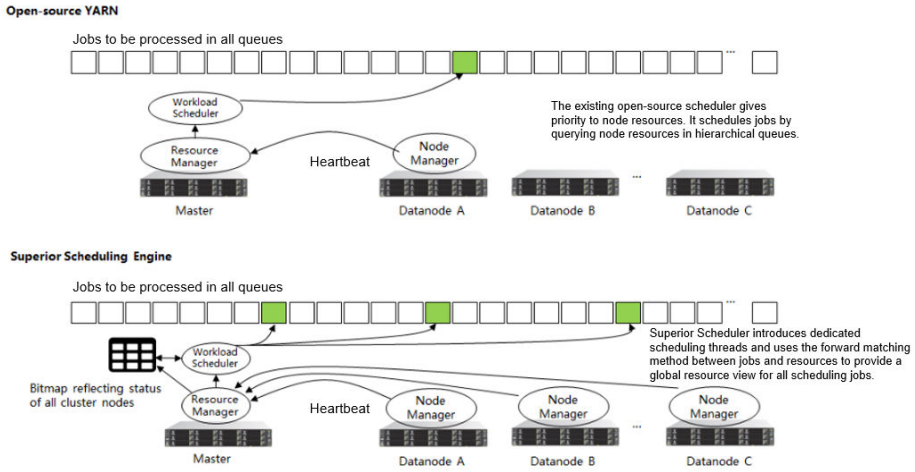
  The scheduling principle of open source schedulers is that resources match jobs based on the heartbeats of computing nodes. Specifically, each computing node periodically sends heartbeat messages to ResourceManager of Yarn to notify the node status and starts the scheduler to assign jobs to the node itself. In this scheduling mechanism, the scheduling period depends on the heartbeat. If the cluster scale increases, bottleneck on system scalability and scheduling performance may occur. In addition, because resources match jobs, the scheduling accuracy of an open source scheduler is limited. For example, data affinity is random and the system does not support load-based scheduling policies. The scheduler may not make an optimal choice due to lack of the global resource view when selecting jobs.

  Superior Scheduler adopts multiple scheduling mechanisms. There are dedicated scheduling threads in Superior Scheduler, separating heartbeats with scheduling and preventing system heartbeat storms. Additionally,

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Superior Scheduler matches jobs with resources, providing each scheduled job with a global resource view and increasing the scheduling accuracy. Compared with the open source scheduler, Superior Scheduler excels in system throughput, resource usage, and data affinity.

Figure 1-142 Comparison of Superior Scheduler with open source schedulers



Apart from the enhanced system throughput and utilization, Superior Scheduler provides following major scheduling features:

- Multiple resource pools

  Multiple resource pools help logically divide cluster resources and share them among multiple tenants or queues. The division of resource pools supports heterogeneous resources. Resource pools can be divided exactly according to requirements on the application resource isolation. You can configure further policies for different queues in a pool.

- Multi-tenant scheduling (**reserve**, **min**, **share**, and **max**) in each resource pool

  Superior Scheduler provides flexible hierarchical multi-tenant scheduling policy. Different policies can be configured for different tenants or queues that can access different resource pools. The following figure lists supported policies:

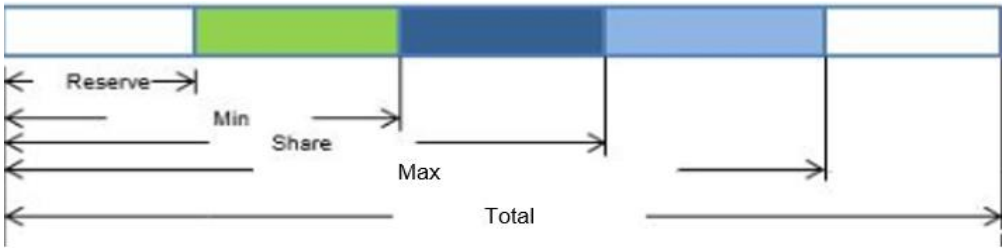
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-31 Policy description

Name	Description
reserve	This policy is used to reserve resources for a tenant. Even though tenant has no jobs available, other tenant cannot use the reserved resource. The value can be a percentage or an absolute value. If both the percentage and absolute value are configured, the percentage is automatically calculated into an absolute value, and the larger value is used. The default **reserve** value is **0**. Compared with the method of specifying a dedicated resource pool and hosts, the **reserve** policy provides a flexible floating reservation function. In addition, because no specific hosts are specified, the data affinity for calculation is improved and the impact by the faulty hosts is avoided.
min	This policy allows preemption of minimum resources. Other tenants can use these resources, but the current tenant has the priority to use them. The value can be a percentage or an absolute value. If both the percentage and absolute value are configured, the percentage is automatically calculated into an absolute value, and the larger value is used. The default value is **0**.
share	This policy is used for shared resources that cannot be preempted. To use these resources, the current tenant needs to wait for other tenants to complete jobs and release resources. The value can be a percentage or an absolute value.
max	This policy is used for the maximum resources that can be utilized. The tenant cannot obtain more resources than the allowed maximum value. The value can be a percentage or an absolute value. If both the percentage and absolute value are configured, the percentage is automatically calculated into an absolute value, and the larger value is used. By default value, there is no restriction on resources.

Figure 1-143 shows the tenant resource allocation policy.

Figure 1-143 Resource scheduling policies



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

In the above figure, **Total** indicates the total number of resources, not the scheduling policy.

Compared with open source schedulers, Superior Scheduler supports both percentage and absolute value of tenants for allocating resources, flexibly addressing resource scheduling requirements of enterprise-level tenants. For example, resources can be allocated according to the absolute value of level-1 tenants, avoiding impact caused by changes of cluster scale. However, resources can be allocated according to the allocation percentage of sub-tenants, improving resource usages in the level-1 tenant.

- Heterogeneous and multi-dimensional resource scheduling

  Superior Scheduler supports following functions except CPU and memory scheduling:

  - Node labels can be used to identify multi-dimensional attributes of different nodes, and can be scheduled based on these labels.
  - Resource pools can be used to group resources of the same type and allocate them to specific tenants or queues.

- Fair scheduling of multiple users in a tenant

  In a leaf tenant, multiple users can use the same queue to submit jobs. Compared with the open source schedulers, Superior Scheduler supports configuring flexible resource sharing policy among different users in a same tenant. For example, VIP users can be configured with higher resource access weight.

- Data locality aware scheduling

  Superior Scheduler adopts the job-to-node scheduling policy. That is, Superior Scheduler attempts to schedule specified jobs between available nodes so that the selected node is suitable for the specified jobs. By doing so, the scheduler will have an overall view of the cluster and data. Localization is ensured if there is an opportunity to place tasks closer to the data. The open source scheduler uses the node-to-job scheduling policy to match the appropriate jobs to a given node.

- Dynamic resource reservation during container scheduling

  In a heterogeneous and diversified computing environment, some containers need more resources or multiple resources. For example, Spark job may require large memory. Such containers may not obtain sufficient resources within a reasonable period when competing with containers requiring fewer resources. Open source schedulers allocate resources to jobs, which may cause unreasonable resource reservation for these jobs. This mechanism leads to the waste of overall system resources. Superior Scheduler differs from open source schedulers in following aspects:

  - Requirement-based matching: Superior Scheduler schedules jobs to nodes and selects appropriate nodes to reserve resources to improve the startup time of containers and avoid waste.
  - Tenant rebalancing: When the reservation logic is enabled, the open source schedulers do not comply with the configured sharing policy. Superior Scheduler uses different methods. In each scheduling period, Superior Scheduler traverses all tenants and attempts to balance resources based on the multi-tenant policy. In addition, Superior

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scheduler attempts to meet all policies (**reserve**, **min**, and **share**) to release reserved resources and direct available resources to other containers that should obtain resources under different tenants.

- Dynamic queue status control (**Open**/**Closed**/**Active**/**Inactive**)

  Multiple queue statuses are supported, helping MRS cluster administrators manage and maintain multiple tenants.

  - Open status (**Open/Closed**): If the status is **Open** by default, applications submitted to the queue are accepted. If the status is **Closed**, no application is accepted.
  - Active status (**Active/Inactive**): If the status is **Active** by default, resources can be scheduled and allocated to applications in the tenant. Resources will not be scheduled to queues in **Inactive** status.

- Application pending reason

  If the application is not started, provide the job pending reasons.

Table 1-32 describes the comparison result of Superior Scheduler and YARN open-source schedulers.

Table 1-32 Comparative analysis

Scheduling	Yarn Open Source Scheduler	Superior Scheduler
Multi-tenant scheduling	In homogeneous clusters, either Capacity Scheduler or Fair Scheduler can be selected and the cluster does not support Fair Scheduler. Capacity Scheduler supports the scheduling by percentage and Fair Scheduler supports the scheduling by absolute value.	• Supports heterogeneous clusters and multiple resource pools. • Supports **reservation** to ensure direct access to resources.
Data locality aware scheduling	The node-to-job scheduling policy reduces the success rate of data localization and potentially affects application execution performance.	The **job-to-node scheduling policy** can aware data location more accurately, and the job hit rate of data localization scheduling is higher.
Balanced scheduling based on load of hosts	Not supported	**Balanced scheduling can be achieved when Superior Scheduler considers the host load and resource allocation during scheduling.**
Fair scheduling of multiple users in a tenant	Not supported	Supports keywords **default** and **others**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scheduling	Yarn Open Source Scheduler	Superior Scheduler
Job waiting reason	Not supported	Job waiting reasons illustrate why a job needs to wait.

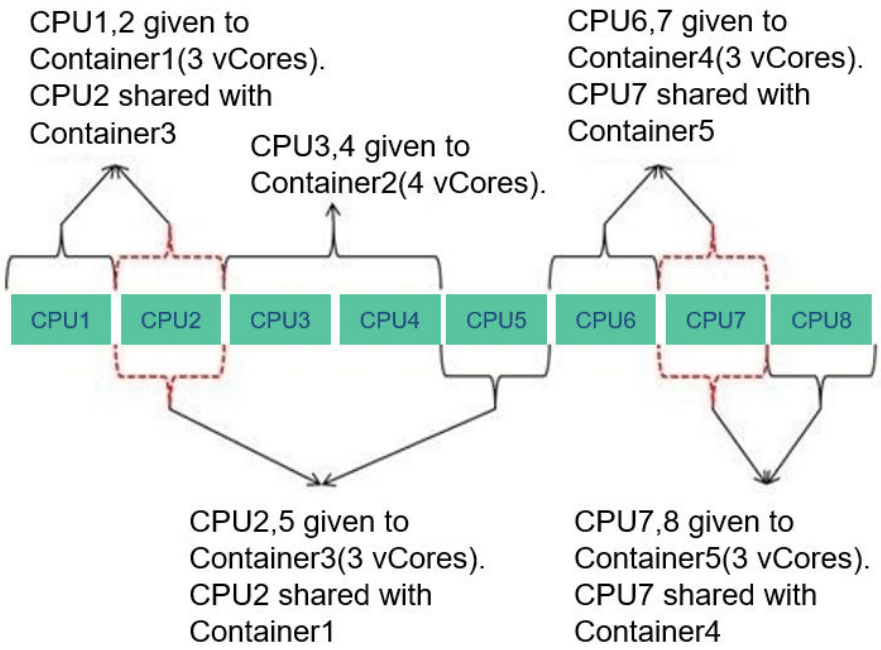
In conclusion, Superior Scheduler is a high-performance scheduler with various scheduling policies and is better than Capacity Scheduler in terms of functionality, performance, resource usage, and scalability.

CPU Hard Isolation

Yarn cannot strictly control the CPU resources used by each container. When the CPU subsystem is used, a container may occupy excessive resources. In this case, CPUset is used to control resource allocation.

To solve this problem, the CPU resources are allocated to each container based on the ratio of virtual cores (vCores) to physical cores. If a container requires an entire physical core, the container has it. If a container needs only some physical cores, several containers may share the same physical core. The following figure shows an example of the CPU quota. The given ratio of vCores to physical cores is 2:1.

Figure 1-144 CPU quota



Enhanced Open Source Feature: Optimizing Restart Performance

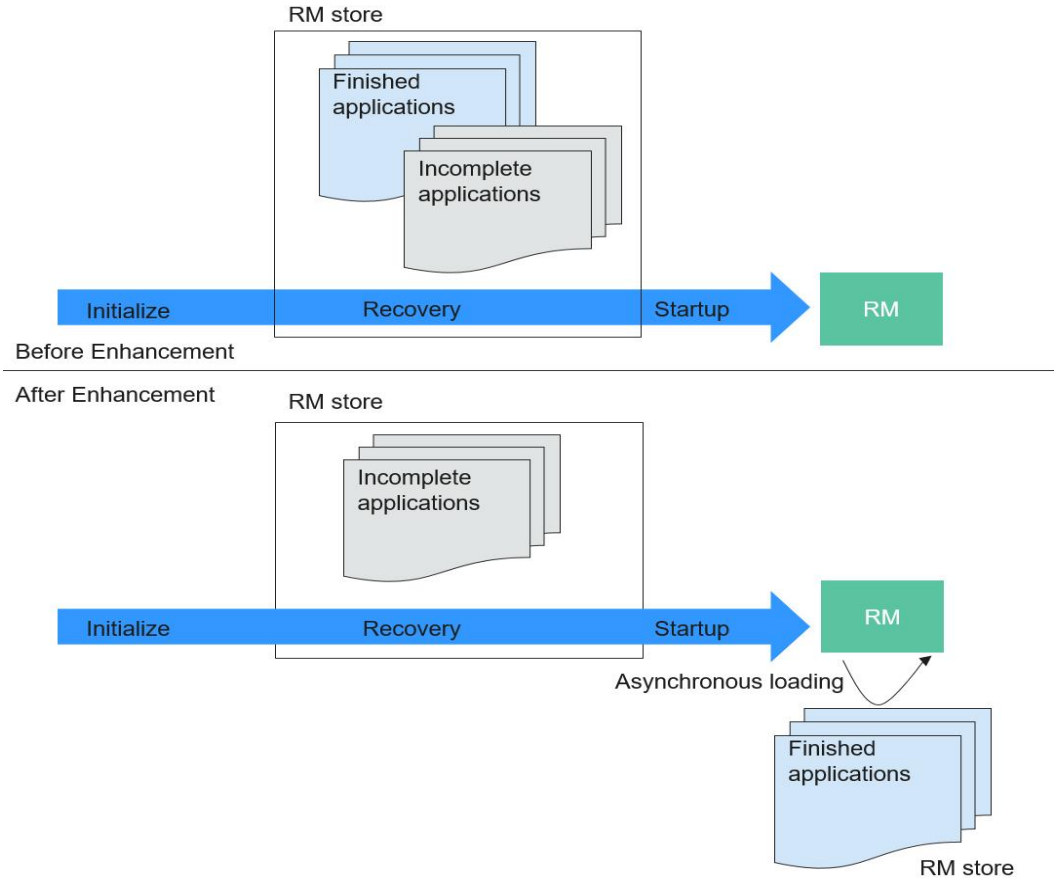
Generally, the recovered ResourceManager can obtain running and completed applications. However, a large number of completed applications may cause

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

problems such as slow startup and long HA switchover/restart time of ResourceManagers.

To speed up the startup, obtain the list of unfinished applications before starting the ResourceManagers. In this case, the completed application continues to be recovered in the background asynchronous thread. The following figure shows how the ResourceManager recovery starts.

Figure 1-145 Starting the ResourceManager recovery



1.6.34 ZooKeeper

1.6.34.1 ZooKeeper Basic Principles

Overview

ZooKeeper is a distributed, highly available coordination service. ZooKeeper is used to provide following functions:

- Prevents the system from SPOFs and provides reliable services for applications.
- Provides distributed coordination services and manages configuration information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Architecture

Nodes in a ZooKeeper cluster have three roles: Leader, Follower, and Observer, as shown in **Figure 1-146**. Generally, an odd number of (2N+1) ZooKeeper services need to be configured in the cluster, and at least (N+1) vote majority is required to successfully perform the write operation.

Figure 1-146 Architecture

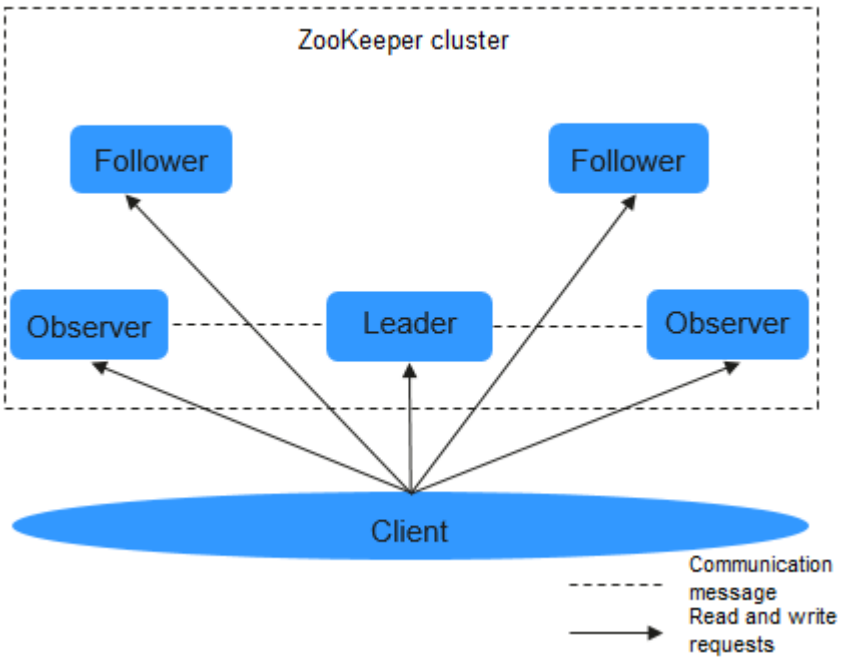


Table 1-33 describes the functions of each module shown in **Figure 1-146**.

Table 1-33 Architecture description

Name	Description
Leader	Only one node serves as the Leader in a ZooKeeper cluster. The Leader, elected by Followers using the ZooKeeper Atomic Broadcast (ZAB) protocol, receives and coordinates all write requests and synchronizes written information to Followers and Observers.
Follower	Follower has two functions: • Prevents SPOFs. A new Leader is elected from Followers when the Leader is faulty. • Processes read requests and interact with the Leader to process write requests.
Observer	The Observer does not take part in voting for election and write requests. It only processes read requests and forwards write requests to the Leader, increasing system processing efficiency.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Description
Client	Reads and writes data from or to the ZooKeeper cluster. For example, HBase can serve as a ZooKeeper client and use the arbitration function of the ZooKeeper cluster to control the active/standby status of HMaster.

If security services are enabled in the cluster, authentication is required during the connection to ZooKeeper. The authentication modes are as follows:

- Keytab mode: You need to obtain a human-machine user from the MRS cluster administrator for MRS console login and authentication, and obtain the Keytab file of the user.
- Ticket mode: Obtain a human-machine user from the MRS cluster administrator for subsequent secure login, enable the renewable and forwardable functions of the Kerberos service, set the ticket update period, and restart Kerberos and related components.

  NOTE

  - By default, the validity period of the user password is 90 days. Therefore, the validity period of the obtained Keytab file is 90 days.
  - The parameters for enabling the renewable and forwardable functions and setting the ticket update interval are on the **System** tab of the Kerberos service configuration page. The ticket update interval can be set to **kdc_renew_lifetime** or **kdc_max_renewable_life** based on the actual situation.

Principles

- **Write Request**
  a. After the Follower or Observer receives a write request, the Follower or Observer sends the request to the Leader.
  b. The Leader coordinates Followers to determine whether to accept the write request by voting.
  c. If more than half of voters return a write success message, the Leader submits the write request and returns a success message. Otherwise, a failure message is returned.
  d. The Follower or Observer returns the processing results.
- **Read-Only Request**
  The client directly reads data from the Leader, Follower, or Observer.

Typical Specifications

Typical Specifications lists the typical specifications of the ZooKeeper service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-34 Typical specifications

Specification	Value	Description
Maximum number of ZooKeeper instances in a cluster	9	Maximum number of ZooKeeper instances.
Maximum number of connections per IP address for each ZooKeeper instance	2000	-
Maximum number of connections for a ZooKeeper instance	20000	-
Maximum number of ZNodes in the case of default configurations	2000000	If there are too many ZNodes, the service will be unstable and the read and write performance of the component deteriorates. In regular service scenarios, it is recommended that there be no more than 2 million ZNodes. If you deployed only ClickHouse and its dependent components in the cluster, there can be no more than 6 million ZNodes.
Size of a single ZNode	4 MB	-

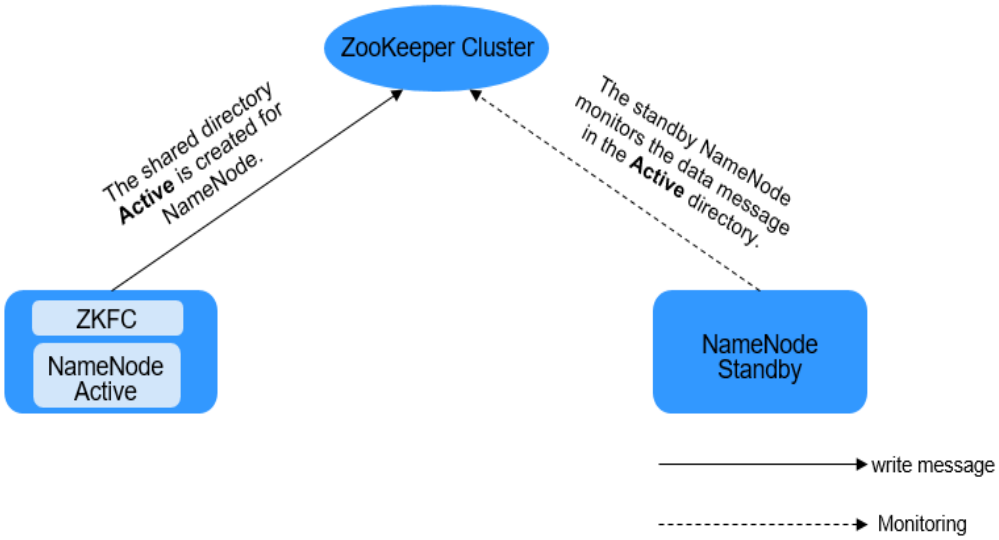
1.6.34.2 Relationships Between ZooKeeper and Other Components

Hadoop Distributed File System (HDFS)

Figure 1-147 shows the relationship between ZooKeeper and HDFS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-147 Relationship between ZooKeeper and HDFS



As the client of a ZooKeeper cluster, ZKFailoverController (ZKFC) monitors the status of NameNode. ZKFC is deployed only in the node where NameNode is deployed, and in both the active and standby HDFS NameNodes.

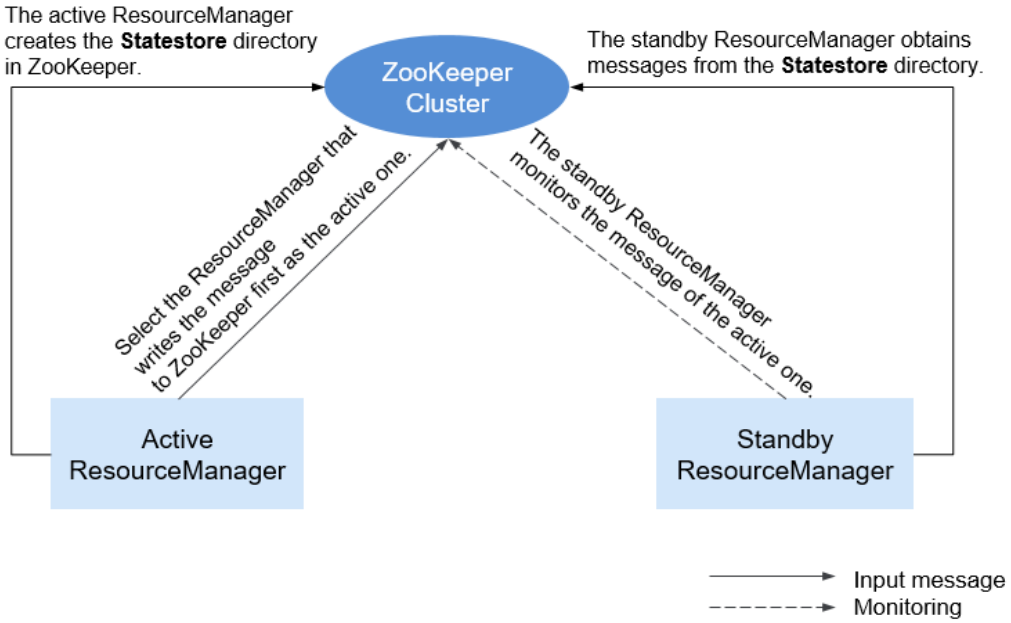
1. The ZKFC connects to ZooKeeper and saves information such as host names to ZooKeeper under the znode directory **/hadoop-ha**. NameNode that creates the directory first is considered as the active node, and the other is the standby node. NameNodes read the NameNode information periodically through ZooKeeper.
2. When the process of the active node ends abnormally, the standby NameNode detects changes in the **/hadoop-ha** directory through ZooKeeper, and then takes over the service of the active NameNode.

YARN

Figure 1-148 shows the relationship between ZooKeeper and YARN.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-148 Relationship between ZooKeeper and YARN

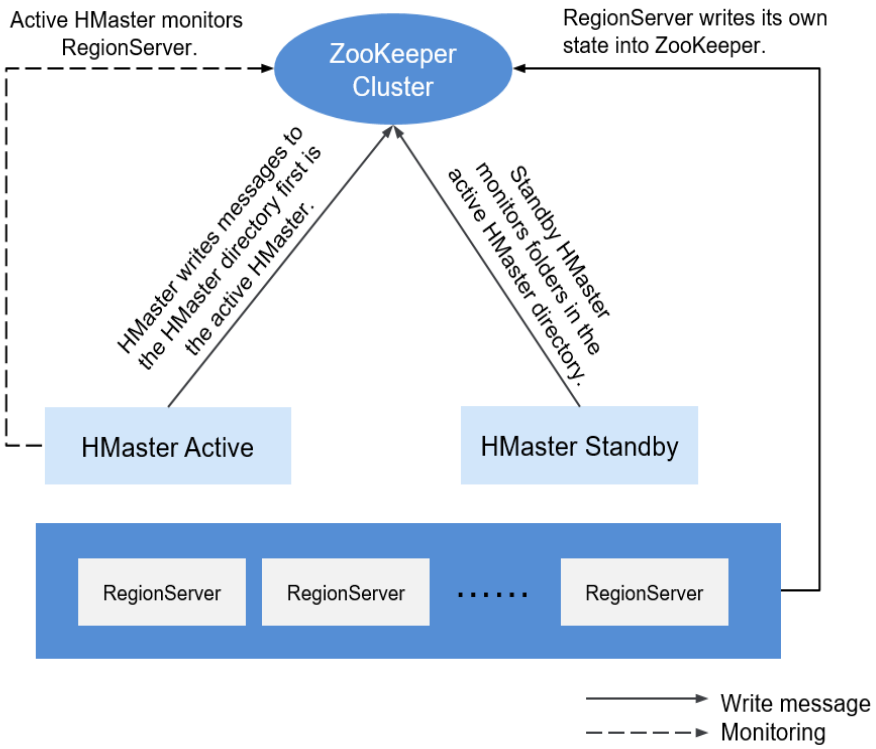


1. When the system is started, ResourceManager attempts to write state information to ZooKeeper. ResourceManager that first writes state information to ZooKeeper is selected as the active ResourceManager, and others are standby ResourceManagers. The standby ResourceManagers periodically monitor active ResourceManager election information in ZooKeeper.
2. The active ResourceManager creates the **Statestore** directory in ZooKeeper to store application information. If the active ResourceManager is faulty, the standby ResourceManager obtains application information from the **Statestore** directory and restores the data.

HBase

Figure 1-149 shows the relationship between ZooKeeper and HBase.

Figure 1-149 Relationship between ZooKeeper and HBase



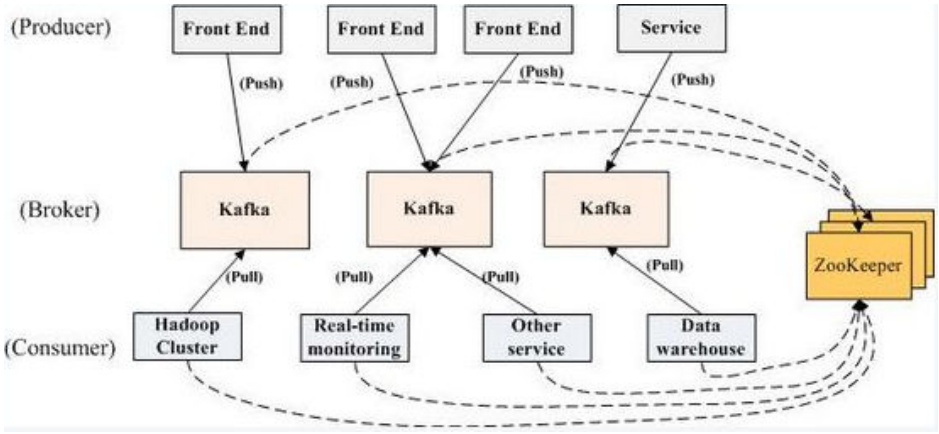
1. RegionServer registers itself to ZooKeeper on Ephemeral node. ZooKeeper stores the HBase information, including the HBase metadata and HMaster addresses.
2. HMaster detects the health status of each RegionServer using ZooKeeper, and monitors them.
3. HBase supports multiple HMaster nodes (like HDFS NameNodes). When the active HMatser is faulty, the standby HMaster obtains the state information about the entire cluster using ZooKeeper. That is, using ZooKeeper can avoid HBase SPOFs.

Kafka

Figure 1-150 shows the relationship between ZooKeeper and Kafka.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 1-150 Relationship between ZooKeeper and Kafka



1. Broker uses ZooKeeper to register broker information and elect a partition leader.
2. The consumer uses ZooKeeper to register consumer information, including the partition list of consumer. In addition, ZooKeeper is used to discover the broker list, establish a socket connection with the partition leader, and obtain messages.

1.6.34.3 ZooKeeper Enhanced Open Source Features

Enhanced Log

In security mode, an ephemeral node is deleted as long as the session that created the node expires. Ephemeral node deletion is recorded in audit logs so that ephemeral node status can be obtained.

Usernames must be added to audit logs for all operations performed on ZooKeeper clients.

On the ZooKeeper client, create a znode, of which the Kerberos principal is **zkcli/hadoop.***<System domain name>***@***<System domain name>*.

For example, open the **<ZOO_LOG_DIR>/zookeeper_audit.log** file. The file content is as follows:

```
2016-12-28 14:17:10,505 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test1?result=success
2016-12-28 14:17:10,530 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test2?result=success
2016-12-28 14:17:10,550 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test3?result=success
2016-12-28 14:17:10,570 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test4?result=success
2016-12-28 14:17:10,592 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test5?result=success
2016-12-28 14:17:10,613 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test6?result=success
2016-12-28 14:17:10,633 | INFO  | CommitProcWorkThread-4 | session=0x12000007553b4903?
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
user=10.177.223.78,zkcli/hadoop.hadoop.com@HADOOP.COM?ip=10.177.223.78?operation=create znode?
target=ZooKeeperServer?znode=/test7?result=success
```

The content shows that logs of the ZooKeeper client user **zkcli/hadoop.hadoop.com@HADOOP.COM** are added to the audit log.

User details in ZooKeeper

In ZooKeeper, different authentication schemes use different credentials as users. Based on the authentication provider requirement, any parameter can be considered as users.

Example:

- **SAMLAuthenticationProvider** uses the client principal as a user.
- **X509AuthenticationProvider** uses the user client certificate as a user.
- **IAuthenticationProvider** uses the client IP address as a user.
- A username can be obtained from the custom authentication provider by implementing the **org.apache.zookeeper.server.auth.ExtAuthenticationProvider.getUserName(String)** method. If the method is not implemented, getting the username from the authentication provider instance will be skipped.

Enhanced Open Source Feature: ZooKeeper SSL Communication (Netty Connection)

The ZooKeeper design contains the Nio package and does not support SSL later than version 3.5. To solve this problem, Netty is added to ZooKeeper. Therefore, if you need to use SSL, enable Netty and set the following parameters on the server and client:

The open source server supports only plain text passwords, which may cause security problems. Therefore, such text passwords are no longer used on the server.

- Client
  a. Set **-Dzookeeper.client.secure** in the **zkCli.sh/zkEnv.sh** file to **true** to use secure communication on the client. Then, the client can connect to the secureClientPort on the server.
  b. Set the following parameters in the **zkCli.sh/zkEnv.sh** file to configure the client environment:

Parameter	Description
-Dzookeeper.clientCnxnSocket	Used for Netty communication between clients. Default value: **org.apache.zookeeper.ClientCnxnSocketNetty**
-Dzookeeper.ssl.keyStore.location	Indicates the path for storing the keystore file.
-Dzookeeper.ssl.keyStore.password	Encrypts a password.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
-Dzookeeper.ssl.trustStore.location	Indicates the path for storing the truststore file.
-Dzookeeper.ssl.trustStore.password	Encrypts a password.
-Dzookeeper.config.crypt.class	Decrypts an encrypted password.
-Dzookeeper.ssl.password.encrypted	Default value: **false** If the keystore and truststore passwords are encrypted, set this parameter to **true**.
-Dzookeeper.ssl.enabled.protocols	Defines the SSL protocols to be enabled for the SSL context.
-Dzookeeper.ssl.exclude.cipher.ext	Defines the list of passwords separated by a comma which should be excluded from the SSL context.

NOTE

The preceding parameters must be set in the **zkCli.sh/zkEnv.sh** file.

- Server
  a. Set **secureClientPort** to **3381** in the **zoo.cfg** file.
  b. Set **zookeeper.serverCnxnFactory** to **org.apache.zookeeper.server.NettyServerCnxnFactory** in the **zoo.cfg** file on the server.
  c. Set the following parameters in the **zoo.cfg** file (in the **zookeeper/conf/zoo.cfg** path) to configure the server environment:

Parameter	Description
ssl.keyStore.location	Path for storing the **keystore.jks** file
ssl.keyStore.password	Encrypts a password.
ssl.trustStore.location	Indicates the path for storing the truststore file.
ssl.trustStore.password	Encrypts a password.
config.crypt.class	Decrypts an encrypted password.
ssl.keyStore.password.encrypted	Default value: **false** If this parameter is set to **true**, the encrypted password can be used.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ssl.trustStore.password.encrypted	Default value: **false** If this parameter is set to **true**, the encrypted password can be used.
ssl.enabled.protocols	Defines the SSL protocols to be enabled for the SSL context.
ssl.exclude.cipher.ext	Defines the list of passwords separated by a comma which should be excluded from the SSL context.

d. Start ZKserver and connect the security client to the security port.

- Credential

  The credential used between client and server in ZooKeeper is **X509AuthenticationProvider**. This credential is initialized using the server certificates specified and trusted by the following parameters:

  – zookeeper.ssl.keyStore.location
  – zookeeper.ssl.keyStore.password
  – zookeeper.ssl.trustStore.location
  – zookeeper.ssl.trustStore.password

  NOTE

  If you do not want to use default mechanism of ZooKeeper, then it can be configured with different trust mechanisms as needed.

1.7 Functions

1.7.1 Cluster Lifecycle Management

MRS supports cluster lifecycle management, including creating and terminating clusters.

- Creating a cluster: After you specify a cluster type, components, number of nodes of each type, VM specifications, AZ, VPC, and authentication information, MRS automatically creates a cluster that meets the configuration requirements. You can run customized scripts in the cluster. In addition, you can create clusters of different types for multiple application scenarios, such as Hadoop analysis clusters, HBase clusters, and Kafka clusters. The big data platform supports heterogeneous cluster deployment. That is, VMs of different specifications can be combined in a cluster based on CPU types, disk capacities, disk types, and memory sizes. Various VM specifications can be mixed in a cluster.
- Terminating a cluster: You can terminate a cluster that is no longer needed (including data and configurations in the cluster). MRS will delete all resources related to the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Billing

The MRS billing is simple and predictable. It supports pay-per-use and yearly/monthly billing modes.

The total price of an MRS cluster will be automatically calculated so you can buy a cluster with a few clicks.

Creating a Cluster

On the MRS management console, you can create an MRS cluster and select a region and cloud resource specifications that are suitable for your company with a few clicks. MRS automatically installs and deploys the enterprise-level big data platform and optimizes parameters based on the selected cluster type, version, and node specifications.

MRS provides you with fully managed big data clusters. When creating a cluster, you can set a VM login mode (password or key pair). You can use all resources of the created MRS cluster.

MRS clusters are classified into analysis, streaming, and hybrid clusters.

- Analysis cluster: is used for offline data analysis and provides Hadoop components.
- Streaming cluster: is used for streaming tasks and provides stream processing components.
- Hybrid cluster: is used for not only offline data analysis but also streaming processing, and provides Hadoop components and stream processing components.
- Custom: You can flexibly combine required components based on service requirements.

MRS cluster nodes are classified into Master, Core, and Task nodes.

- Master nodes are management nodes in a cluster. The Master process, Manager, and databases of the distributed system are deployed on these nodes. Master nodes cannot be expanded. The processing capability of Master nodes determines the upper limit of the management capability of the entire cluster. MRS supports scale-up of Master node specifications to provide support for management of a larger cluster.
- Core node: used for both storage and computing and can be scaled in or out. Since Core nodes bear data storage, there are many restrictions on scale-in to prevent data loss and auto scaling cannot be performed.
- Task node: used only for computing only and can be scaled in or out. Task nodes bear only computing tasks. Therefore, auto scaling can be performed.

You can create a cluster in two modes: custom configuration and quick configuration.

- **Custom Config**: On the **Custom Config** page, you can flexibly configure cluster parameters based on application scenarios, such as ECS specifications to better suit your service requirements.
- **Quick Config**: On the **Quick Config** page, you can quickly create a cluster based on application scenarios, improving cluster configuration efficiency.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Terminating a Cluster

MRS allows you to terminate a cluster when it is no longer needed. After the cluster is terminated, all cloud resources used by the cluster will be released. Before terminating a cluster, you are advised to migrate or back up data. Terminate the cluster only when no service is running in the cluster or the cluster is abnormal and cannot provide services based on O&M analysis.

If data is stored on EVS disks or pass-through disks in a big data cluster, the data will be deleted after the cluster is terminated. Therefore, exercise caution when terminating a cluster.

1.7.2 Enterprise Project Management

An enterprise project is a cloud resource management mode. Enterprise Management provides users with comprehensive management of cloud-based resources, personnel, permissions, and finances. Common management consoles are oriented to the control and configuration of individual cloud products. The Enterprise Management console, in contrast, is more focused on resource management. It is designed to help enterprises manage cloud-based resources, personnel, permissions, and finances, in a hierarchical management manner, such as management of companies, departments, and projects.

MRS allows users who have enabled Enterprise Project Management Service (EPS) to configure enterprise projects for a cluster during cluster creation and use EPS to manage MRS resources by group.

- The users can manage multiple resources by group.
- The users can view resource information and expenditure details of enterprise projects.
- The users can control access permissions at the enterprise project level.
- The users can view detailed financial information by enterprise project, including orders, expenditure summary, and expenditure details.

NOTE

If the MRS cluster and VPC are not in the same enterprise project, you need to add the VPC view permission in the IAM view to view VPC and cluster information.

1.7.3 Job Management

The MRS job management function provides an entry for you to submit jobs in a cluster, including MapReduce, Spark, HQL, and SparkSQL jobs.

MRS works with DataArts Studio to provide a one-stop big data collaboration development environment and fully-managed big data scheduling capabilities, helping you effortlessly build big data processing centers.

DataArts Studio allows you to develop and debug MRS HQL/SparkSQL scripts online and develop MRS jobs by performing drag-and-drop operations to migrate and integrate data between MRS and over 20 heterogeneous data sources. Powerful job scheduling and flexible monitoring and alarming help you easily manage data and job O&M.

You can create the following types of jobs on the console in an MRS cluster:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- MapReduce can quickly process large-scale data in parallel. It is a distributed data processing model and execution environment. MRS supports the submission of MapReduce JAR programs.
- Spark is a distributed in-memory computing framework. MRS supports SparkSubmit, Spark Script, and Spark SQL jobs.
  - SparkSubmit: You can submit Spark JAR and Spark Python programs, execute the Spark Application, and compute and process user data.
  - SparkScript: You can submit SparkScript scripts and batch execute Spark SQL statements.
  - Spark SQL: You can use Spark SQL statements (similar to SQL statements) to query and analyze user data in real time.
- Hive is an open-source data warehouse based on Hadoop. MRS allows you to submit HiveScript scripts and directly execute Hive SQL statements.
- Flink is a distributed big data processing engine that can perform stateful computations over both unbounded and bounded data streams.
- Hadoop Streaming works similarly to a standard Hadoop job, where you can define the input and output HDFS paths, as well as the mapper and reducer executable programs.

1.7.4 Metadata Management

MRS provides multiple metadata storage methods. When deploying Hive and Ranger during MRS cluster creation, select one of the following storage modes as required:

- **Local**: Metadata is stored in the local GaussDB of a cluster. When the cluster is deleted, the metadata is also deleted. To retain the metadata, manually back up the metadata in the database in advance.
- **External data connection**: After the cluster is created, you can select **RDS PostgreSQL database** or **RDS MySQL database** that is associated with the same VPC and subnet as the current cluster. Metadata is stored in the associated database and is not deleted when the current cluster is deleted. Multiple MRS clusters can share the same metadata.

NOTE

Hive in MRS 1.9.*x* or later allows you to specify a metadata storage method.

1.7.5 Managing Multi-Tenancy Resources

Feature Introduction

Modern enterprises' data clusters are developing towards centralization and cloudification. Enterprise-class big data clusters must meet the following requirements:

- Carry data of different types and formats and run jobs and applications of different types (analysis, query, and stream processing).
- Isolate data of a user from that of another user who has high requirements on data security, such as a bank or government institute.

The preceding requirements bring the following challenges to the big data cluster:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Proper allocation and scheduling of resources to ensure stable operating of applications and jobs
- Strict access control to ensure data and service security

Multi-tenant isolates the resources of a big data cluster into resource sets. Users can lease desired resource sets to run applications and jobs and store data. In a big data cluster, multiple resource sets can be deployed to meet diverse requirements of multiple users.

The MRS big data cluster provides a complete enterprise-class big data multi-tenant solution. Multi-tenant is a collection of multiple resources (each resource set is a tenant) in an MRS big data cluster. It can allocate and schedule resources, including computing and storage resources.

Advantages

- Proper resource configuration and isolation

  The resources of a tenant are isolated from those of another tenant. The resource use of a tenant does not affect other tenants. This mechanism ensures that each tenant can configure resources based on service requirements, improving resource utilization.
- Resource consumption measurement and statistics

  Tenants are system resource applicants and consumers. System resources are planned and allocated based on tenants. Resource consumption by tenants can be measured and recorded.
- Ensured data security and access security

  In multi-tenant scenarios, the data of each tenant is stored separately to ensure data security. The access to tenants' resources is controlled to ensure access security.

Enhanced Schedulers

Schedulers are divided into the open source Capacity scheduler and proprietary Superior scheduler.

To meet enterprise requirements and tackle challenges facing the Yarn community in scheduling, develops the Superior scheduler. In addition to inheriting the advantages of the Capacity scheduler and Fair scheduler, this scheduler is enhanced in the following aspects:

- Enhanced resource sharing policy

  The Superior scheduler supports queue hierarchy. It integrates the functions of open source schedulers and shares resources based on configurable policies. In terms of instances, MRS cluster administrators can use the Superior scheduler to configure an absolute value or percentage policy for queue resources. The resource sharing policy of the Superior scheduler enhances the label scheduling policy of Yarn as a resource pool feature. The nodes in the Yarn cluster can be grouped based on the capacity or service type to ensure that queues can more efficiently utilize resources.
- Tenant-based resource reservation policy

  Resources required by tenants must be ensured for running critical tasks. The Superior scheduler builds a mechanism to support the resource reservation

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

policy. By doing so, reserved resources can be allocated to the tasks run by the tenant queues in a timely manner to ensure proper task execution.

- Fair sharing among tenants and resource pool users

  The Superior scheduler allows shared resources to be configured for users in a queue. Each tenant may have users with different weights. Heavily weighted users may require more shared resources.

- Ensured scheduling performance in a big cluster

  The Superior scheduler receives heartbeats from each NodeManager and saves resource information in memory, which enables the scheduler to control cluster resource usage globally. The Superior scheduler uses the push scheduling model, which makes the scheduling more precise and efficient and remarkably improves cluster resource utilization. Additionally, the Superior scheduler delivers excellent performance when the interval between NodeManager heartbeats is long and prevents heartbeat storms in big clusters.

- Priority policy

  If the minimum resource requirement of a service cannot be met after the service obtains all available resources, a preemption occurs. The preemption function is disabled by default.

1.7.6 Easy Access to Web UIs of Components

Each big data component has its own web interface for system management. Unfortunately, network isolation prevents easy access to these interfaces.

For instance, to access the HDFS web UI, you must first create an ECS instance for remote login. This adds complexity to accessing the UI.

With MRS, you can securely access component web UIs through a pre-configured EIP-based channel, eliminating the need to manually bind an EIP or perform login, security group setup, and public IP acquisition tasks.

You can quickly access the web UIs of components in the analysis cluster, such as Hadoop, Spark, HBase, Hue, and streaming cluster, on Manager.

1.7.7 Cluster Management

1.7.7.1 Cluster Scaling

The processing capability of a big data cluster can be horizontally expanded by adding nodes. If the cluster scale does not meet service requirements, you can manually scale out or scale in the cluster. MRS intelligently selects the node with the least load or the minimum amount of data to be migrated for scale-in. The node to be scaled in will not receive new tasks, and continues to execute the existing tasks. At the same time, MRS copies its data to other nodes and the node is decommissioned. If the tasks on the node cannot be completed after a long time, MRS migrates the tasks to other nodes, minimizing the impact on cluster services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scaling Out a Cluster

Currently, you can add Core or Task nodes to scale out a cluster to handle peak service loads.

During and after the scale-out of the MRS cluster nodes, services in the existing cluster are not affected. If data skew occurs after the scale-out of data components, you need to manually balance data.

Scaling In a Cluster

You can remove Core or Task nodes to scale in a cluster based on service requirements. This helps MRS deliver better storage and computing capabilities at lower O&M costs. After you scale in an MRS cluster, MRS automatically selects nodes that can be scaled in based on the type of services installed on the nodes.

During the scale-in of Core nodes, data on the original nodes is migrated. If the data location is cached, the client automatically updates the location information, which may affect the latency. Node scale-in may affect the response duration of the first access to some HBase on HDFS data. You can restart HBase or disable or enable related tables to avoid this problem.

Task nodes do not store cluster data. They are compute nodes and do not involve migration of data on the nodes.

1.7.7.2 Creating Task Nodes

Feature Introduction

Task nodes can be created and used for computing only. They do not store persistent data and are the basis for implementing auto scaling.

Customer Benefits

When MRS is used only as a computing resource, Task nodes can be used to reduce costs and facilitate cluster node scaling, flexibly meeting users' requirements for increasing or decreasing cluster computing capabilities.

Application Scenarios

If the cluster data volume is stable but the processing demand increase greatly, you can add Task nodes to the cluster to handle more service requests.

- The number of temporary services is increased, for example, report processing at the end of the year.
- A large number of tasks, for example, emergent analysis tasks, need to be processed in a short period.

1.7.7.3 Auto Scaling

Feature Introduction

More and more enterprises use technologies such as Spark and Hive to analyze data. Processing a large amount of data consumes huge resources and costs

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

much. Typically, enterprises regularly analyze data in a fixed period of time every day rather than all day long. To meet enterprises' requirements, MRS provides the auto scaling function to apply for extra resources during peak hours and release resources during off-peak hours. This enables users to use resources on demand and focus on core business at lower costs.

In big data applications, especially in periodic data analysis and processing scenarios, cluster computing resources need to be dynamically adjusted based on service data changes to meet service requirements. The auto scaling function of MRS enables clusters to be elastically scaled out or in based on cluster loads. In addition, if the data volume changes regularly and you want to scale out or in a cluster before the data volume changes, you can use the MRS resource plan feature.

MRS supports two types of auto scaling policies: auto scaling rules and resource plans

- Auto scaling rules: You can increase or decrease Task nodes based on real-time cluster loads. Auto scaling will be triggered when the data volume changes but there may be some delay.
- Resource plans: If the data volume changes periodically, you can create resource plans to resize the cluster before the data volume changes, thereby avoiding a delay in increasing or decreasing resources.

Both auto scaling rules and resource plans can trigger auto scaling. You can configure both of them or configure one of them. Configuring both resource plans and auto scaling rules improves the cluster node scalability to cope with occasionally unexpected data volume peaks.

In some service scenarios, resources need to be reallocated or service logic needs to be modified after cluster scale-out or scale-in. If you manually scale a cluster, you can log in to cluster nodes to reallocate resources or modify service logic. If you use auto scaling, MRS enables you to customize automation scripts for resource reallocation and service logic modification. Automation scripts can be executed before and after auto scaling and automatically adapt to service load changes, all of which eliminates manual operations. In addition, automation scripts can be fully customized and executed at various moments, which can meet your personalized requirements and improve auto scaling flexibility.

Customer Benefits

MRS auto scaling provides the following benefits:

- Reducing costs

  Enterprises may not analyze data in batches all the time but perform a batch analysis task in a specified period of time, for example, 03:00 a.m. The task may take only two hours.

  The auto scaling function allows you to add nodes for batch analysis and automatically releases the nodes after completion of the analysis, minimizing costs.

- Meeting instant query requirements

  Enterprises usually encounter instant analysis tasks, for example, data reports for supporting enterprise decision-making. As a result, resource consumption increases sharply in a short period of time. With the auto scaling function,

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

compute resources can be added for emergent big data analysis, avoiding service breakdown due to insufficient compute nodes. After the service spike ends and you do not need additional resources, MRS schedules and performs a scale-in.

- Focusing on core business

  It is difficult for developers to determine resource consumption on the big data secondary development platform because of complex query analysis conditions (such as global sorting, filtering, and merging) and data complexity, for example, uncertainty of incremental data. As a result, estimating the computing volume is difficult. MRS's auto scaling function enables developers to focus on service development without the need for resource estimation.

1.7.7.4 Isolating Nodes

When detecting that a host is abnormal or faulty and cannot provide services or affects cluster performance, you can exclude the host from the available nodes in the cluster temporarily so that the client can access other available nodes.

In scenarios where patches are to be installed in an MRS cluster, you can also exclude a specified node from patch installation.

NOTE

Only non-management nodes can be isolated.

When a host is isolated, all role instances on it are automatically stopped, and you cannot start, stop or configure them.

When a host is isolated, you can no longer collect or view statistics for its monitoring status and hardware/instance indicators.

1.7.7.5 Scaling Up Master Node Specifications

MRS provides Manager for managing clusters and services in the clusters, such as NameNodes of HDFS, ResourceManagers of Yarn, and Manager management services of MRS, are deployed on the Master node of the clusters.

With the rollout of new services, a cluster scale increases continuously, and Master nodes bear more and more loads. Enterprise users are faced with the problem that CPU loads are too high and memory usage exceeds the threshold. Generally, in an on-premises big data cluster, you need to migrate data and purchase hardware with advanced configurations to scale up the Master node specifications. MRS leverages the advantages of cloud services to enable you to scale up Master node specifications in one click. During the scale-up, the active/standby HA mode of the Master nodes ensures that existing services are not interrupted.

1.7.7.6 Managing Node Labels

Tags help you identify and manage your clusters and nodes more easily. By associating with Tag Management Service (TMS), MRS allows users with a large number of cloud resources to tag cloud resources, quickly search for cloud resources with the same tag attribute, and perform unified management operations such as review, modification, and deletion, facilitating unified management of big data clusters and other cloud resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can add a maximum of 10 tags to a cluster when creating the cluster or add them on the details page of the created cluster.

You can add tags to a node only after the node is created. You can add a maximum of 10 tags to a node.

1.7.7.7 Node Bootstrap Actions

Feature Introduction

MRS provides standard elastic big data clusters on the cloud. Big data components, such as Hadoop and Spark, can be installed and deployed. Currently, standard cloud big data clusters cannot meet all user requirements, for example, in the following scenarios:

- Common operating system configurations cannot meet data processing requirements, for example, increasing the maximum number of system connections.
- Software tools or running environments need to be installed, for example, Gradle and dependency R language package.
- Big data component packages need to be modified based on service requirements, for example, modifying the Hadoop or Spark installation package.
- Other big data components that are not supported by MRS need to be installed.

To meet the preceding customization requirements, you can manually perform operations on the existing and newly added nodes. The overall process is complex and error-prone. In addition, manual operations cannot be traced, and data cannot be processed immediately after creating a cluster based on your demand.

Therefore, MRS supports custom bootstrap actions that enable you to run scripts on a specified node before or after a cluster component is started. You can run bootstrap actions to install third-party software that is not supported by MRS, modify the cluster running environment, and perform other customizations. If you choose to run bootstrap actions when expanding a cluster, the bootstrap actions will be run on the newly added nodes in the same way. MRS runs the script you specify as user **root**. You can run the **su - *xxx*** command in the script to switch the user.

Customer Benefits

You can use the custom bootstrap actions to flexibly and easily configure your dedicated clusters and customize software installation.

1.7.7.8 Cluster Status Notification

Feature Introduction

The following operations are often performed during the running of a big data cluster:

- Big data clusters often change, for example, cluster scale-out and scale-in.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- When a service data volume changes abruptly, auto scaling will be triggered.
- After related services are stopped, a big data cluster needs to be stopped.

To immediately notify you of successful operations, cluster unavailability, and node faults, MRS uses Simple Message Notification (SMN) to send notifications to you through SMS and emails, facilitating maintenance.

Customer Benefits

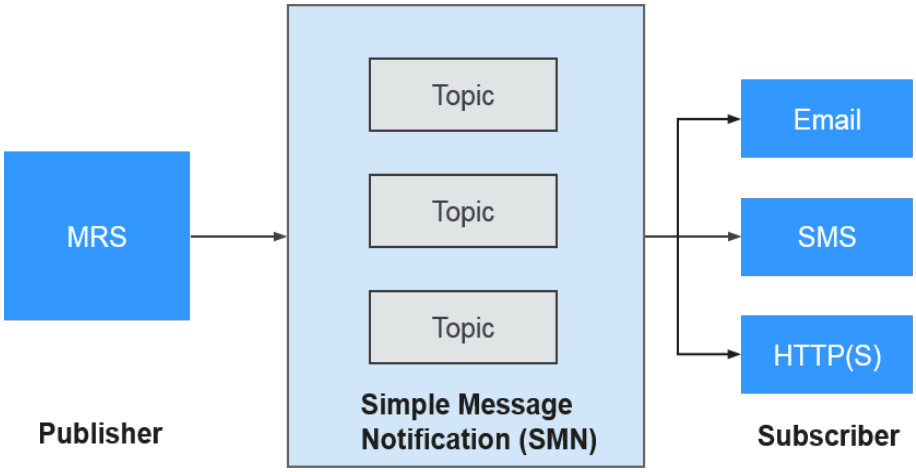
After configuring SMN, you can receive MRS cluster health status, updates, and component alarms through SMS or emails in real time. MRS sends real-time monitoring and alarm notification to help you easily perform O&M and efficiently deploy big data services.

Feature Description

MRS uses SMN to provide one-to-multiple message subscription and notification over a variety of protocols.

You can create a topic and configure topic policies to control publisher and subscriber permissions on the topic. MRS sends cluster messages to the topic to which you have permission to publish messages. Then, all subscribers who subscribe to the topic can receive cluster updates and component alarms through SMS and emails.

Figure 1-151 MRS cluster status notification



1.7.8 Cluster O&M

Cluster resources provided by MRS are completely owned by users who thereby can use various methods to maintain cluster running.

Manager

MRS manages and analyzes massive amounts of data and quickly mines valuable data from structured and unstructured data. In view of the complex structure of open-source components and time- and labor-consuming installation, configuration, and management, Manager provides a unified enterprise-level big data cluster management platform with the following features:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Cluster monitoring: enables you to quickly learn the running status of hosts and services.
- Graphical metric monitoring and customization: enable you to quickly obtain key information about the system.
- Service property configuration: allows you to configure service properties based on the performance requirements of your services.
- Cluster, service, and role instance operations: allow you to start or stop services and clusters with just a few clicks.

Alarm Management

MRS can monitor big data clusters in real time and identify system health status based on alarms and events. In addition, MRS allows you to customize monitoring and alarm thresholds to focus on the health status of each metric. When monitoring data reaches the alarm threshold, the system triggers an alarm.

MRS can connect to Simple Message Notification (SMN) to send alarm information to users by SMS messages or emails. For details, see **Cluster Status Notification**.

Audit

All operations on MRS Manager are recorded for post-event tracing, fault locating, and responsibility division in security events.

Patch Management

MRS supports cluster patching operations and will release patches for open source big data components in a timely manner. On the MRS cluster management page, you can view patch release information related to running clusters, including the detailed description of the resolved issues and impacts. You can determine whether to install a patch based on the service running status. One-click patch installation involves no manual intervention, and will not cause service interruption through rolling installation, ensuring long-term availability of the clusters.

MRS can display the detailed patch installation process, and supports patch uninstallation and rollback.

Cluster Health Checks

MRS provides automatic inspection on system running environments for you to check and audit system running health status in one click, ensuring proper system running and lowering system operation and maintenance costs. After viewing inspection results, you can export reports for archiving and fault analysis.

Performing a Rolling Restart of Services

To apply configuration changes to a big data component, you must restart it. However, the common restart will restart all services or instances, which may cause service interruption.

To minimize or eliminate the impact on services during a component restart, you can perform rolling restarts to restart components or instances in batches. For

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

instances in active/standby mode, the standby instance is restarted first, followed by the active instance.

Rolling restart takes longer than normal restart.

O&M Support

Cluster resources provided by MRS belong to users. Generally, when O&M personnel's support is required for troubleshooting of a cluster, O&M personnel cannot directly access the cluster. To better serve customers, MRS provides the following two methods to improve communication efficiency during fault locating:

- Log sharing: You can initiate log sharing on the MRS management console to share a specified log scope with O&M personnel, so that O&M personnel can locate faults without accessing the cluster.
- O&M authorization: If a problem occurs when you use an MRS cluster, you can initiate O&M authorization on the MRS management console. O&M personnel can help you quickly locate the problem, and you can revoke the authorization at any time.

1.7.9 MRS Security Hardening

MRS is a platform for massive data management and analysis and has high security. MRS protects user data and service running from the following aspects:

- Network isolation

  The entire system is deployed in a VPC on the cloud to provide an isolated network environment and ensure service and management security of the cluster. By combining the subnet division, route control, and security group functions of VPC, MRS provides a secure and reliable isolated network environment.
- Resource isolation

  MRS supports resource deployment and isolation of physical resources in dedicated zones. You can flexibly combine computing and storage resources, such as dedicated computing resources + shared storage resources, shared computing resources + dedicated storage resources, and dedicated computing resources + dedicated storage resources.
- Host security

  MRS can be integrated with security services, including Vulnerability Scan Service (VSS), Host Security Service (HSS), Web Application Firewall (WAF), Cloud Bastion Host (CBH), and Web Tamper Protection (WTP). The following measures are provided to improve security of the OS and ports:
  - Security hardening of OS kernels
  - OS permission control
  - OS port management
- Application security

  The following measures are used to ensure normal running of big data services:
  - Identification and authentication
  - Web application security

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Access control
  - Audit security
  - Password security
- Data security

  The following measures are provided to ensure the confidentiality, integrity, and availability of massive amounts of user data:

  - Disaster recovery: MRS supports data backup to OBS and cross-region high reliability.
  - Backup: MRS supports backup of DBService, NameNode, and LDAP metadata and backup of HDFS and HBase service data.
- Data integrity

  Data is verified to ensure its integrity during storage and transmission.

  - CRC32C is used by default to verify the correctness of user data stored in HDFS.
  - DataNodes of HDFS store the verified data. If the data transmitted from a client is abnormal (incomplete), DataNodes report the abnormality to the client, and the client rewrites the data.
  - The client checks data integrity when reading data from a DataNode. If the data is incomplete, the client will read data from another DataNode.
- Data confidentiality

  Based on Apache Hadoop, the distributed file system of MRS supports encrypted storage of files to prevent sensitive data from being stored in plaintext, improving data security. Applications need only to encrypt specified sensitive data. Services are not affected during the encryption process. Based on file system data encryption, Hive provides table-level encryption and HBase provides column family-level encryption. Sensitive data can be encrypted and stored after you specify an encryption algorithm during table creation.

  Encrypted storage and access control of data are used to ensure user data security.

  - HBase can compress service data before storing it to HDFS. You can configure AES and SM4 (also called SMS4) algorithms for encrypted storage.
  - All the components allow access permissions to be set for local data directories. Unauthorized users are not allowed to access data.
  - All cluster user information is stored in ciphertext.
- Security authentication
  - Uses a unified user- and role-based authentication system as well as an account- and role-based access control (RBAC) model to centrally control user permissions and batch manage user authorization.
  - Employs Lightweight Directory Access Protocol (LDAP) as an account management system and performs the Kerberos authentication on accounts.
  - Provides the single sign-on (SSO) function that centrally manages and authenticates MRS system and component users.
  - Audits users who have logged in to Manager.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1.7.10 MRS Reliability Enhancement

Based on Apache Hadoop open source software, MRS optimizes and improves the reliability and performance of main service components.

System Reliability

- HA for all management nodes

  In the Hadoop open source version, data and compute nodes are managed in a distributed system, in which a single point of failure (SPOF) does not affect the operation of the entire system. However, a SPOF may occur on management nodes running in centralized mode, which becomes the weakness of the overall system reliability.

  MRS provides similar double-node mechanisms for all management nodes of the service components, such as Manager, HDFS NameNodes, HiveServers, HBase HMasters, Yarn ResourceManagers, KerberosServers, and LdapServers. All of them are deployed in active/standby mode or configured with load sharing, effectively preventing SPOFs from affecting system reliability.

- Reliability guarantee in case of exceptions

  By reliability analysis, the following measures to handle software and hardware exceptions are provided to improve the system reliability:

  - After power supply is restored, services are running properly regardless of a power failure of a single node or the whole cluster, ensuring data reliability in case of unexpected power failures. Key data will not be lost unless the hard disk is damaged.
  - Health status checks and fault handling of the hard disk do not affect services.
  - The file system faults can be automatically handled, and affected services can be automatically restored.
  - The process and node faults can be automatically handled, and affected services can be automatically restored.
  - The network faults can be automatically handled, and affected services can be automatically restored.

- Data backup and restoration

  MRS provides full backup, incremental backup, and restoration functions based on service requirements, preventing the impact of data loss and damages on services and ensuring fast system restoration in case of exceptions.

  - Automatic backup

    MRS provides automatic backup for data on Manager. Based on the customized backup policy, data on clusters, including LdapServer and DBService data, can be automatically backed up.

  - Manual backup

    You can also manually back up data of the cluster management system before the capacity expansion and patch installation to recover the cluster management system functions upon faults.

    To improve the system reliability, data on Manager and HBase is backed up to a third-party server manually.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Node Reliability

- OS health status monitoring

  MRS periodically collects OS hardware resource usage data, including usage of CPUs, memory, hard disks, and network resources.

- Process health status monitoring

  MRS checks the status of service instances and health indicators of service instance processes, enabling you to know the health status of processes in a timely manner.

- Automatic disk troubleshooting

  MRS is enhanced based on the open source version. It can monitor the status of hardware and file systems on all nodes. If an exception occurs, the corresponding partitions will be removed from the storage pool. If a disk is faulty and replaced, a new hard disk will be added for running services. In this case, maintenance operations are simplified. Replacement of faulty disks can be completed online. In addition, users can set hot backup disks to reduce the faulty disk restoration time and improve the system reliability.

- LVM configuration for node disks

  MRS allows you to configure Logic Volume Management (LVM) to plan multiple disks as a logical volume group. Configuring LVM can avoid uneven usage of disks. It is especially important to ensure even usage of disks on components that can use multiple disk capabilities, such as HDFS and Kafka. In addition, LVM supports disk capacity expansion without re-attaching, preventing service interruption.

Data Reliability

MRS can use the anti-affinity node groups provided by ECS and the rack awareness capability of Hadoop to redundantly distribute data to multiple physical host machines, preventing data loss caused by physical hardware failures.

1.8 Constraints

Before using MRS, ensure that you have read and understood the following restrictions.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Cluster Creation Constraints

Table 1-35 Constraints on creating an MRS cluster

Constraint	Description
Network requirement	• MRS clusters must be created in VPC subnets. • When you create an MRS cluster, you can select **Auto create** from **Security Group** to create a security group or select an existing security group. • To prevent illegal access, only assign access permission for security groups used by MRS where necessary. • To ensure that the cluster runs properly, do not use the following IP addresses for cluster nodes or areas that the cluster can access: 10.10.10.10, 10.10.10.11, and 1.1.1.1.
Browser	You are advised to use a recommended browser to log in to the MRS management page. • Google Chrome: 36.0 or later • Microsoft Edge is updated with the Windows operating system.
Data storage	• The cluster nodes store only users' service data. Non-service data can be stored in the OBS or other ECS nodes. • The cluster nodes only run MRS cluster programs. Other client applications or user service programs are deployed on separate ECS nodes. • Disk resources for each cluster node should be planned according to service demands. For services that require significant data storage, allocate additional EVS disks or expand storage capacity to avoid storage shortages impacting node performance. • The capacity (including storage and computing capabilities) of an MRS cluster can be expanded by adding core or task nodes.
Password requirement	Keep the initial password for logging in to the master node properly because MRS does not save it. Use a complex password to avoid malicious attacks.
Technical support	• If a cluster exception occurs when no incorrect operations have been performed, contact technical support engineers. They will ask you for your password and then perform troubleshooting.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS Cluster Running Constraints

Table 1-36 MRS cluster running constraints

Item	Description
Node management	• If a Master node in an MRS cluster is shut down and the cluster is still used to execute jobs or modify component configurations, you must start the stopped Master node before stopping other nodes. Otherwise, data may be lost due to an active/standby switchover. • If all nodes in an MRS cluster have been stopped, start them in the reverse order of node shutdown.
Resources scheduling	When the MRS cluster is running, the scheduler can be switched between Capacity and Superior but the configuration synchronization cannot be ensured. Configure synchronization again based on the new scheduler if necessary.
Storage-compute decoupling	To delete a component or cluster connected to OBS (including storage-compute decoupling and cold-hot data separation scenarios), you must also delete the service data on OBS.

Components	• Do not enable the MOB feature for the HBase service in the MRS cluster. Using this feature may lead to table data read failure and JVM crash.
	For existing HBase tables, run the following command in **hbase shell** to check whether the table description contains the keyword **MOB**. If it is contained, contact O&M engineers to set the table non-MOB.
	desc '*Table name*'
	For example, if the value of **IS_MOB** is **true** in the following command output, the HBase MOB feature is enabled:
	`hbase:009:0> desc 't3'`
	`t3`
	`COLUMN FAMILIES DESCRIPTION`
	`{NAME => 'd', MOB_THRESHOLD => '102400', VERSIONS => '1',`
	`KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL =>`
	`'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOO`
	`MFILTER => 'ROW', IN_MEMORY => 'false', IS_MOB => 'true', COMPRESSION =>`
	`'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}`
	• Spark and other engines in the MRS cluster are forbidden to execute self-read and self-write SQL statements.
	For example, the following SQL statement overwrites the same data with the data it reads:
	`insert overwrite table test select * from test;`
	`insert overwrite table test select a.* from test a join test2 b on a.name=b.name;`

Forbidden and High-Risk Operations on Clusters and Components

To avoid service disruptions due to unstable clusters, do not perform the following actions during the use of MRS clusters and components.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-37 Forbidden operations on MRS clusters

Operation	Risk
• Shutting down, restarting, or deleting MRS cluster nodes on the ECS console, changing or reinstalling their OS, or modifying their specifications. • Deleting the existing processes, applications, and files from cluster nodes.	The cluster will function abnormally, affecting its usability.
After the MRS cluster is created, do not delete or modify the used security group.	A cluster exception may occur.
Deleting ZooKeeper data directories.	Component metadata stored in ZooKeeper is lost, including ClickHouse, HDFS, Yarn, HBase, and Hive. Upper-layer components may experience performance issues or errors.
Switching between active and standby JDBCServer nodes frequently.	Services may be interrupted.
Deleting Phoenix system tables or table data (SYSTEM.CATALOG, SYSTEM.STATS, SYSTEM.SEQUENCE, and SYSTEM. FUNCTION).	Service operations will fail.
Modifying data in the Hive metabase (**hivemeta** database).	Hive data parse errors may occur. As a result, Hive cannot provide services.
Performing **INSERT** or **UPDATE** operations on Hive metadata tables.	Modifying Hive metadata may cause Hive data parsing errors and affect Hive services.
Changing permission on the Hive private file directory **hdfs:///tmp/hive-scratch**.	Hive services may be unavailable.
Modifying **broker.id** in the Kafka configuration file.	Data on the node may become invalid.
Modifying the host names of nodes.	Instances and upper-layer components on the host cannot provide services properly. The fault cannot be rectified.
Using private images.	The cluster will function abnormally, affecting its usability.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	Risk
Modifying the default DNS configurations of MRS cluster nodes	The nodes may fail to resolve domain names of other servers, causing communication interruption and cluster exceptions.

Table 1-38 lists high-risk operations during maintenance of each component in the MRS cluster.

Table 1-38 High-risk operations on MRS clusters

Operation Type	Operation	Risk	Preventive Measure
Cluster operations	Modify the file directory or file permissions of user **omm** without permission.	This operation will lead to MRS service unavailability.	None
	Bind an EIP.	This operation exposes the Master node hosting MRS Manager of the cluster to the public network, increasing the risk of network attacks from the Internet.	Ensure that the bound EIP is a trusted public IP address.
	Enable security group rules for port 22 of a cluster.	This operation increases the risk of exploiting vulnerability of port 22.	Configure a security group rule for port 22 to allow only trusted IP addresses to access the port. You are not advised to configure the inbound rule to allow 0.0.0.0 to access the port.
	Delete a cluster or cluster data.	This operation will cause data loss.	Before deleting the data, confirm the necessity of the operation and ensure that the data has been backed up.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Scale in a cluster.	This operation will cause data loss.	Before scaling in the cluster, confirm the necessity of the operation and ensure that the data has been backed up.
	Detach or format a data disk.	This operation will cause data loss.	Before performing this operation, confirm the necessity of the operation and ensure that the data has been backed up.
Manager operations	Change the OMS password.	This operation will restart all processes of OMS, which has adverse impact on cluster maintenance and management.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Import the certificate.	This operation will restart OMS processes and the entire cluster, which has adverse impact on cluster maintenance and management and services.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Restore the OMS.	This operation will restart Manager and the entire cluster, affecting management, maintenance, and services of the cluster.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Change log levels.	If the log level is changed to **DEBUG**, Manager responds slowly.	Before the modification, confirm the necessity of the operation and change it back to the default log level in time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Restart both upper-layer and lower-layer services.	This operation will interrupt the upper-layer service, affecting the management, maintenance, and services of the cluster.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Modify the OLDAP port.	This operation will restart the LdapServer and Kerberos services and all associated services, affecting service running.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Delete the **supergroup** user group bound to a user.	This operation decreases user rights, affecting service access.	Before the change, confirm the rights to be added. Ensure that the required rights have been added before deleting the **supergroup** rights to which the user is bound, ensuring service continuity.
	Restart or stop services.	Services will be interrupted during the restart. If you select and restart the upper-layer service, the upper-layer services that depend on the service will be interrupted.	Confirm the necessity of restarting the system before the operation.
	Change the default SSH port.	Changing the default port (22) leads to incorrect health check results, including the inspection of node mutual trust and **omm**/**ommdba** user password expiration.	Before performing this operation, restore the SSH port to the default value.
ClickHouse	Delete the ClickHouse data directory.	This operation may cause service information loss.	Do not delete data directories.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Remove ClickHouseServer instances.	This operation leads to incorrect topology information for the logical cluster. To avoid this, all ClickHouse Server instance nodes within the same shard must be scaled in and decommissioned at the same time. Before performing this operation, check the database and data table information of each node in the logical cluster and perform scale-in pre-analysis to ensure that data can be migrated during scale-in and decommissioning, preventing data loss.	Before scale-in, collect information in advance to learn the status of the ClickHouse logical cluster and instance nodes. Check the ClickHouse logical cluster topology information, database and data table information in each ClickHouseServer instance, and data volume.
	Add ClickHouseServer instances.	Before performing this operation, verify if a database or table with the matching name exists on the new node; otherwise, subsequent operations such as data migration, balancing, scaling down, and decommissioning will fail.	Before scale-out, confirm the function and purpose of new ClickHouseServer instances and determine whether to create related databases and data tables.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Decommission ClickHouseServer instances.	This operation leads to incorrect topology information for the logical cluster. To avoid this, all ClickHouse Server instance nodes within the same shard must be decommissioneda at the same time. Before performing this operation, check the database and data table information of each node in the logical cluster and perform pre-analysis to ensure that data can be migrated during decommissioning, preventing data loss.	Before decommissioning, collect information in advance to learn the status of the ClickHouse logical cluster and instance nodes.
	Recommission ClickHouseServer instances.	When performing this operation, you must select all nodes in the original shard. Otherwise, the topology information of the logical cluster is incorrect.	Before recommissioning, you need to confirm the home information about the shards of the node to be recommissioned.
	Modify data directory content (file and folder creation).	This operation may cause ClickHouse instance faults on the node.	Do not create or modify files or folders in the data directories.
	Start or stop basic components independently.	This operation has adverse impact on the basic functions of some services. As a result, service failures occur.	Do not start or stop ZooKeeper, Kerberos, and LDAP basic components independently. Select related services when performing this operation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Use the built-in **default** user and **clickhouse** user.	The passwords of the built-in **default** user and **clickhouse** user are uncertain, which may cause production service loss.	Create and use a service user.
DBService	Change the DBService password.	The services need to be restarted for the password change to take effect. The services are unavailable during the restart.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Restore DBService data.	After the data is restored, the data generated after the data backup and before the data restoration is lost. After the data is restored, the configurations of the components that depend on DBService may expire and these components need to be restarted.	Perform the operation only when it is necessary, and ensure that there is no other maintenance and management operations when the operation is performed.
	Perform active/standby DBService switchover.	During the DBServer switchover, DBService is unavailable.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
Flink	Change the Flink log level.	If the log level is modified to DEBUG, the task running performance is affected.	Before the modification, confirm the necessity of the operation and change it back to the default log level in time.
	Modify Flink file permissions.	This operation may cause task execution failures.	Before the modification, confirm that the modification is needed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
Flume	Modify the Flume instance start parameter **GC_OPTS**.	Flume cannot start.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
HBase	Modify encryption configuration. • hbase.regionserver.wal.encryption • hbase.crypto.keyprovider.parameters.uri • hbase.crypto.keyprovider.parameters.encryptedtext	The HBase service fails to be started.	Strictly follow the prompt information when modifying related configuration items, which are associated. Ensure that new values are valid.
	Disable or switch the encryption algorithm when encryption is enabled.	The encryption function will be disabled when **hbase.regionserver.wal.encryption** is set to **false**. The algorithm can be switched to AES or SM4 (also called SMS4). These operations may cause service startup failure and data loss.	When HFile and WAL are encrypted using an encryption algorithm and a table is created, do not disable encryption or switch the encryption algorithm. If an encryption table (ENCRYPTION=>AES/SMS4) is not created, you can only switch the encryption algorithm.
	Modify HBase instance start parameter **GC_OPTS** and **HBASE_HEAPSIZE**.	The HBase service fails to be started.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid. **GC_OPTS** does not conflict with **HBASE_HEAPSIZE**.
	Use **OfflineMetaRepair** tool.	The HBase service fails to be started.	This tool can be used only when HBase is offline and cannot be used in data migration scenarios.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
HDFS	Change HDFS NameNode data storage directory **dfs.namenode.name.dir** and DataNode data configuration directory **dfs.datanode.data.dir**.	The HDFS service fails to be started.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Use the **-delete** parameter when you run the **hadoop distcp** command.	During DistCP copying, files that do not exist in the source cluster but exist in the destination cluster are deleted from the destination cluster.	When using DistCP, determine whether to retain the redundant files in the destination cluster. Exercise caution when using the **-delete** parameter. After DistCP copying is complete, check whether the data in the destination cluster is retained or deleted according to the parameter settings.
	Modify the HDFS instance start parameter **GC_OPTS**, **HADOOP_HEAPSIZE**, and **GC_PROFILE**.	The HDFS service fails to be started.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid. **GC_OPTS** does not conflict with **HADOOP_HEAPSIZE**.
	Change the value of **dfs.replication** (number of HDFS replicas) from **3** to **1**.	• The storage reliability deteriorates. If the disk becomes faulty, data will be lost. • NameNode fails to be restarted, and the HDFS service is unavailable.	When modifying related configuration items, check the parameter description carefully. Ensure that there are more than two replicas for data storage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Change the encryption mode (**hadoop.rpc.protection**) of the remote procedure call (RPC) channel in each module of Hadoop.	The HDFS service is faulty and abnormal.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
Hive	Modify the Hive instance start parameter **GC_OPTS**.	This operation may cause Hive instance start failures.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Delete all MetaStore instances.	Hive metadata is lost, and Hive cannot provide services.	Do not perform this operation unless ensure that Hive table information can be discarded.
	Delete or modify files corresponding to Hive tables over HDFS interfaces or HBase interfaces.	This operation may cause Hive service data loss or tampering.	Do not perform this operation unless ensure that the data can be discarded or that the operation meets service requirements.
	Delete or modify files corresponding to Hive tables or directory access permission over HDFS interfaces or HBase interfaces.	This operation may cause related services unavailable.	Do not perform this operation.
	Delete or modify **hdfs:///apps/templeton/hive-3.1.0.tar.gz** using HDFS interfaces.	WebHCat fails to perform services due to this operation.	Do not perform this operation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Export table data to overwrite the data at the local. For example, export the data of **t1** to **/opt/dir**. **insert overwrite local directory '/opt/dir' select * from t1;**	This operation will delete target directories. Incorrect setting may cause software or OS startup failures.	Ensure that the path where the data is written does not contain any files or do not use the key word overwrite in the command.
	Direct different databases, tables, or partition files to the same path, for example, default warehouse path **/user/hive/warehouse**.	The creation operation may cause disordered data. After a database, table, or partition is deleted, other object data will be lost.	Do not perform this operation.
IoTDB	Delete data directories.	This operation may cause service information loss.	Do not delete data directories.
	Modify data directory content (file and folder creation).	This operation may cause IoTDB instance faults on the node.	Do not create or modify files or folders in the data directories.
	Start or stop basic components independently.	This operation has adverse impact on the basic functions of some services. As a result, service failures occur.	Do not start or stop basic components such as Kerberos and LDAP independently. Select related services when performing this operation.
Kafka	Delete topics	This operation will delete the existing schema and data.	Use Kerberos authentication to ensure that authenticated users have operation permissions. Ensure that topic names are correct.
	Delete the Kafka data directory.	This operation may cause service information loss.	Do not delete data directories.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Modify data directory content (file and folder creation).	This operation may cause the Broker instance of the node faults.	Do not create or modify files or folders in the data directories.
	Modify the disk auto-adaptation function using the **disk.adapter.enable** parameter.	This operation adjusts the topic data retention period when the disk usage reaches the threshold. Historical data that does not fall within the storage retention may be deleted.	If the storage period of a topic cannot be adjusted, specify the topic in the **disk.adapter.topic.blacklist** parameter and observe the data storage period on the Kafka topic monitoring page.
	Modify data directory **log.dirs** configuration.	If the configuration is incorrect, the process may be faulty.	Ensure that the data directory is empty and the permission is obtained.
	Reduce the capacity of the Kafka cluster.	This operation will reduce the number of data replicas of some topics and may cause topic access failures.	Transfer data replicas before capacity reduction.
	Start or stop basic components independently.	This operation has adverse impact on the basic functions of some services and may cause service failure.	Do not start or stop ZooKeeper, Kerberos, and LDAP basic components independently. Select related services when performing this operation.
	Delete or modify metadata.	Modifying or deleting Kafka metadata on ZooKeeper may cause topic or Kafka service unavailability.	Do not delete or modify Kafka metadata stored on ZooKeeper.
	Delete metadata backup files.	After Kafka metadata backup files are modified and used to restore Kafka metadata, Kafka topics or the Kafka service may be unavailable.	Do not delete Kafka metadata backup files.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
KrbServer	Modify the **KADMIN_PORT** parameter of KrbServer.	After this parameter is modified, if the KrbServer service and its associated services are not restarted in a timely manner, the configuration of KrbClient in the cluster is abnormal and the service running is affected.	After this parameter is modified, restart the KrbServer service and all its associated services.
	Modify the **kdc_ports** parameter of KrbServer.	After this parameter is modified, if the KrbServer service and its associated services are not restarted in a timely manner, the configuration of KrbClient in the cluster is abnormal and the service running is affected.	After this parameter is modified, restart the KrbServer service and all its associated services.
	Modify the **KPASSWD_PORT** parameter of KrbServer.	After this parameter is modified, if the KrbServer service and its associated services are not restarted in a timely manner, the configuration of KrbClient in the cluster is abnormal and the service running is affected.	After this parameter is modified, restart the KrbServer service and all its associated services.
	Modify the domain name of Manager.	After the domain name is modified, if the KrbServer service and its associated services are not restarted in a timely manner, the configuration of KrbClient in the cluster is abnormal and the service running is affected.	After this parameter is modified, restart the KrbServer service and all its associated services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Configure cross-cluster mutual trust relationships.	This operation will restart the KrbServer service and all associated services, affecting the management and maintenance and services of the cluster.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
LdapServer	Modify the **LDAP_SERVER_PORT** parameter of LdapServer.	After this parameter is modified, if the LdapServer service and its associated services are not restarted in a timely manner, the configuration of LdapClient in the cluster is abnormal and the service running is affected.	After this parameter is modified, restart the LdapServer service and all its associated services.
	Restore LdapServer data.	This operation will restart the Manager and the entire cluster, which has adverse impact on cluster maintenance and management and services.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Replace the Node where LdapServer is located.	This operation will interrupt services deployed on the node. If the node is a management node, the operation will restart all OMS processes, affecting cluster management and maintenance.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.
	Change the password of LdapServer.	The LdapServer and Kerberos services need to be restarted during the password change, affecting the management, maintenance, and services of the cluster.	Before performing the operation, ensure that the operation is necessary, and that no other management and maintenance operations are performed at the same time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Restart the node where LdapServer is located.	Restarting the node without stopping the LdapServer service may cause LdapServer data damage.	Restore the LdapServer node using the backup data.
Loader	Change the floating IP address of a Loader instance.	This operation causes Loader start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Modify the Loader instance start parameter **LOADER_GC_OPTS**.	This operation causes Loader start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Clear a table when adding data to HBase.	This operation will clear the original data in the target table.	Ensure that the target table can be cleared before the operation.
Spark	Modify the **spark.yarn.queue** configuration item.	This operation causes Spark start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Modifying the **spark.driver.extraJavaOptions** configuration item	This operation causes Spark start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Modify the **spark.yarn.cluster.driver.extraJavaOptions** configuration item.	This operation causes Spark start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Modify the **spark.eventLog.dir** configuration item.	This operation causes Spark start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Modify the **SPARK_DAEMON_JAVA_OPTS** configuration item.	This operation causes Spark start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Delete all JobHistory instances.	The event logs of historical applications are lost.	Retain at least one JobHistory instance and check whether historical application information can be queried in JobHistory.
	Delete the **spark-archive** file from in the HDFS or modify the file.	JDBCServer fails to be started and service functions are abnormal.	After **/user/spark2x/jars/*XXX*/spark-archive-2x.zip** or **/user/spark/jars/*XXX*/spark-archive.zip** is deleted, wait for 10 to 15 minutes. The **.zip** package is automatically restored.
Storm	Modify plug-in related configuration items, including the following: ● storm.scheduler ● nimbus.authorizer ● storm.thrift.transport ● nimbus.blobstore.class ● nimbus.topology.validator ● storm.principal.tolocal	This operation causes Storm start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that the class names exist and are valid.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
	Modify Storm instance **GC_OPTS** startup parameters, including: • NIMBUS_GC_OPTS • SUPERVISOR_GC_OPTS • UI_GC_OPTS • LOGVIEWER_GC_OPTS	This operation causes Storm start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Modify the user resource pool configuration parameter **resource.aware.scheduler.user.pools**.	This operation causes Storm execution failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that resources allocated to each user are appropriate and valid.
	Modify the Storm data directory.	This operation causes services abnormal and unavailable.	Do not manually change data directories.
	Delete or modify Storm metadata.	Deleting Nimbus metadata will cause service exceptions and loss of running services.	Do not manually delete Nimbus metadata files.
	Modify Storm file permission.	If permissions on the metadata and log directories are incorrectly modified, service exceptions may occur.	Do not modify file permissions.
	Delete a Storm topology.	Topologies in use will be deleted.	Delete topologies only when necessary.
Yarn	Delete or modify data directories **yarn.nodemanager.local-dirs** and **yarn.nodemanager.log-dirs**.	This operation causes service information loss.	Do not delete data directories.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation	Risk	Preventive Measure
ZooKeeper	Delete or change ZooKeeper data directories.	This operation causes service information loss.	Strictly follow the capacity expansion guide to change the ZooKeeper data directories.
	Modify the ZooKeeper instance start parameter **GC_OPTS**.	This operation causes ZooKeeper start failure.	Strictly follow the prompt information when modifying related configuration items. Ensure that new values are valid.
	Modify the znode ACL information in ZooKeeper.	If znode permission is modified in ZooKeeper, other users may have no permission to access the znode and some system functions are abnormal.	Ensure that the modified ACL information does not affect the normal use of ZooKeeper by other components.

1.9 Permission Management

If you need to assign different permissions to personnel in your enterprise to access your MRS resources, Identity and Access Management (IAM) is a good choice for fine-grained permissions management. IAM provides identity authentication, permissions management, and access control, helping you to securely access your cloud resources. If your cloud account does not require IAM for permissions management, you can skip this section.

IAM is a free service. You only pay for the resources in your account.

With IAM, you can control access to specific cloud resources. For example, if you want some software developers in your enterprise to use MRS resources but do not want them to delete MRS resources or perform any other high-risk operations, you can grant permission to use MRS resources but not permission to delete them.

IAM supports role/policy-based authorization and identity policy-based authorization.

The following table describes the differences between these two authorization models.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-39 Differences between role/policy-based and identity policy-based authorization

Authorization Model	Core Relationship	Permissions	Authorization Method	Description
Role/Policy	Roles	• System-defined roles • System-defined policies • Custom policies	Granting roles or policies to principals	To authorize a user, you need to add it to a user group first and then specify the scope of authorization. It is hard to provide fine-grained permissions control using authorization by user groups and a limited number of condition keys. This method is suitable for small- and medium-sized enterprises.
Identity policy	Policies	• System-defined policies • Custom policies	• Granting policies to principals • Attaching policies to principals	You can authorize a user by attaching an identity policy to it. User-specific authorization and a variety of key conditions allow for more fine-grained permissions control. However, this model can be hard to set up. It requires a certain amount of expertise and is suitable for medium- and large-sized enterprises.

Policies and actions in the two authorization models are not interoperable. You are advised to use the ABAC authorization model. For details about system-defined permissions, see **Role/Policy-based Permissions Management** and **Identity Policy-based Permissions Management**.

Role/Policy-based Permissions Management

MRS supports role/policy-based authorization. New IAM users do not have any permissions assigned by default. You need to first add them to one or more groups and attach policies or roles to these groups. The users then inherit permissions from the groups and can perform specified operations on cloud services based on the permissions they have been assigned.

MRS is a project-level service deployed for specific regions. When you set **Scope** to **Region-specific projects** and select the specified projects in the specified regions , the users only have permissions for MRS clusters in the selected projects. If you set

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scope to **All resources**, the users have permissions for MRS clusters in all region-specific projects. When accessing MRS, the users need to switch to the authorized region.

Table 1 lists all the system-defined permissions for MRS. System-defined policies in role/policy-based authorization are not interoperable with those in identity policy-based authorization.

Table 1-40 System-defined permissions for MRS

Role/Policy Name	Description	Type	Dependency
MRS FullAccess	Administrator permissions for MRS. Users with these permissions can perform all operations on MRS resources.	System-defined policy	None
MRS CommonOperations	Common user permissions for MRS. Users with these permissions can use MRS but cannot add or delete resources.	System-defined policy	None
MRS ReadOnlyAccess	Read-only permission for MRS. Users with these permissions can only view MRS resources.	System-defined policy	None
MRS Administrator	Permissions include: • Performing all operations on MRS • Permissions of the **Tenant Guest** and **Server Administrator** policies, which must also be granted to the users **Tenant Guest**: a global role, which must be assigned in the global project. **Server Administrator**: a project-level role, which must be assigned in the same project.	System-defined role	The **Tenant Guest** and **Server Administrator** roles need to be assigned in the same project.

Table 1-41 lists the common operations supported by system-defined permissions for MRS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-41 Common operations supported by system-defined permissions

Operation	MRS FullAccess	MRS CommonOperations	MRS ReadOnlyAccess	MRS Administrator
Creating a cluster	√	x	x	√
Resizing a cluster	√	x	x	√
Upgrading node specifications	√	x	x	√
Deleting a cluster	√	x	x	√
Querying cluster details	√	√	√	√
Querying a cluster list	√	√	√	√
Configuring an auto scaling rule	√	x	x	√
Querying a host list	√	√	√	√
Querying operation logs	√	√	√	√
Creating and executing a job	√	√	x	√
Stopping a job	√	√	x	√
Deleting a job	√	√	x	√
Deleting jobs in batches	√	√	x	√
Querying job details	√	√	√	√

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	MRS FullAccess	MRS CommonOperations	MRS ReadOnlyAccess	MRS Administrator
Querying a job list	√	√	√	√
Creating a folder	√	√	x	√
Deleting a file	√	√	x	√
Querying a file list	√	√	√	√
Operating cluster tags in batches	√	√	x	√
Creating a cluster tag	√	√	x	√
Deleting a cluster tag	√	√	x	√
Querying a resource list by tag	√	√	√	√
Querying cluster tags	√	√	√	√
Accessing Manager	√	√	x	√
Querying a patch list	√	√	√	√
Installing a patch	√	√	x	√
Uninstalling a patch	√	√	x	√
Authorizing O&M channels	√	√	x	√
Sharing O&M channel logs	√	√	x	√

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	MRS FullAccess	MRS CommonOperations	MRS ReadOnlyAccess	MRS Administrator
Querying an alarm list	√	√	√	√
Subscribing to alarm notifications	√	√	x	√
Submitting an SQL statement	√	√	x	√
Querying SQL query results	√	√	x	√
Canceling an SQL execution job	√	√	x	√

Identity Policy-based Permissions Management

MRS supports identity policy-based authorization. **Table 1** lists all the system-defined policies for MRS with identity policy-based authorization. System-defined policies in identity policy-based authorization are not interoperable with those in role/policy-based authorization.

Table 1-42 System-defined policies for MRS

Policy Name	Description	Type
MRSFullAccessPolicy	Administrator permissions for MRS. Users with these permissions can operate and use MRS clusters.	System-defined identity policy
MRSReadOnlyAccess-Policy	Read-only permissions for MRS. Users with these permissions can only view MRS resources.	System-defined identity policy
autoSystemPolicy-Name5Mrs	Automated test permissions	System-defined identity policy

Table 1-43 lists the common operations supported by system-defined policies for MRS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 1-43 Common operations supported by system-defined policies

Operation	MRSFullAccessPolicy	MRSReadOnlyAccessPolicy
Creating a cluster	√	x
Resizing a cluster	√	x
Upgrading node specifications	√	x
Deleting a cluster	√	x
Querying cluster details	√	√
Querying a cluster list	√	√
Configuring an auto scaling rule	√	x
Querying a host list	√	√
Querying operation logs	√	√
Creating and executing a job	√	x
Stopping a job	√	x
Deleting a job	√	x
Deleting jobs in batches	√	x
Querying job details	√	√
Querying a job list	√	√
Creating a folder	√	x
Deleting a file	√	x
Querying a file list	√	√
Operating cluster tags in batches	√	x
Creating a cluster tag	√	x
Deleting a cluster tag	√	x
Querying a resource list by tag	√	√
Querying cluster tags	√	√
Accessing Manager	√	x

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	MRSFullAccessPolicy	MRSReadOnlyAccessPolicy
Querying a patch list	√	√
Installing a patch	√	x
Uninstalling a patch	√	x
Authorizing O&M channels	√	x
Sharing O&M channel logs	√	x
Querying an alarm list	√	√
Subscribing to alarm notifications	√	x
Submitting an SQL statement	√	x
Querying SQL query results	√	x
Canceling an SQL execution job	√	x

1.10 Related Services

Relationships with Other Services

Table 1-44 Relationships with other services

Service	Relationships
Virtual Private Cloud (VPC)	MRS clusters are created in the subnets of a VPC. VPCs provide a secure, isolated, and logical network environment for your MRS clusters.
Object Storage Service (OBS)	OBS stores the following user data: • MRS job input data, such as user programs and data files • MRS job output data, such as result files and log files of jobs In MRS clusters, HDFS, Hive, MapReduce, YARN, Spark, Flume, and Loader can import or export data from OBS. MRS uses the parallel file system of OBS to provide services.
Elastic Cloud Server (ECS)	MRS uses ECSs as cluster nodes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Relationships
Relational Database Service (RDS)	RDS stores MRS system running data, including MRS cluster metadata.
Identity and Access Management (IAM)	IAM provides authentication for MRS.
Simple Message Notification (SMN)	MRS uses SMN to provide one-to-multiple message subscription and notification over a variety of protocols.
Cloud Trace Service (CTS)	CTS provides you with operation records of MRS resource operation requests and request results for querying, auditing, and backtracking.

Table 1-45 MRS operations recorded by CTS

Operation	Resource Type	Trace Name
Creating a cluster	cluster_mrs	createCluster
Deleting a cluster	cluster_mrs	deleteCluster
Expanding a cluster	cluster_mrs	scaleOutCluster
Shrinking a cluster	cluster_mrs	scaleInCluster
Stopping jobs	job_mrs	batch_delete_job
Stopping a job	job_mrs	kill_job
Submitting a job	job_mrs	submit_job
Authorizing for a cluster node	cluster_mrs	authorizedOM

After you enable CTS, the system starts recording operations on cloud resources. You can view operation records of the last 7 days on the CTS management console. For details, see the *Cloud Trace Service (CTS) User Guide*.

1.11 Quota Description

Available resource quotas are configured for each user account in an environment to prevent resource abuse.

The following lists the resources used by MRS. Quotas are managed by each basic service. If you need to increase quotas, contact technical support of the corresponding service.

- ECS
- BMS

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- VPC
- EVS
- Image Management Service (IMS)
- OBS
- EIP
- SMN
- IAM

1.12 Common Concepts

HBase Table

An HBase table is a three-dimensional map comprised of one or more columns or rows of data.

Column

Column is a dimension of an HBase table. The column name is in the format of *<family>*:*<label>*, where *<family>* and *<label>* can be any combination of characters. An HBase table consists of a set of column families. Each column in the HBase table belongs to a column family.

Column Family

A column family is a collection of columns stored in the HBase schema. To create columns, you must create a column family first. A column family organizes data with the same property in HBase. Each row of data in the same column family is stored on the same server. Each column family can be one attribute, such as compressed packages, timestamps, and data block cache.

MemStore

MemStore is a core of HBase storage. When the amount of data stored in WAL reaches the upper limit, the data is loaded to MemStore for sorting and storage.

RegionServer

RegionServer is a service running on each DataNode in the HBase cluster. It is responsible for serving and managing regions, uploading the load information of regions, and managing distributed master nodes.

Timestamp

A timestamp is a 64-bit integer used to index different versions of the same data. A timestamp can be automatically assigned by HBase when data is written or assigned by users.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Store

Store is a core of HBase storage. A Store hosts one MemStore and multiple StoreFiles. A Store corresponds to a column family of a table in a region.

Index

An index is a data structure that improves the efficiency of data retrieval in a database table. One or more columns in a database table can be used for fast random retrieval of data and efficient access to ordered records.

Coprocessor

A coprocessor is an interface provided by HBase for implementing calculation logic on RegionServer. Coprocessors are classified into system coprocessors and table coprocessors. The former can import all data tables on RegionServer, and the latter can process a specified table.

Block Pool

A block pool is a collection of blocks that belong to a single namespace. DataNodes store blocks from all block pools in a cluster. Each block pool is managed independently, which allows a namespace to generate an ID for a new block without relying on other namespaces. If one NameNode is invalid, the DataNode can still provide services for other NameNodes in the cluster.

DataNode

A DataNode is a worker node in the HDFS cluster. Scheduled by the client or NameNode, DataNodes store and retrieve data and periodically report file blocks to NameNodes.

File Block

A file block is the minimum logical unit stored in the HDFS. Each HDFS file is stored in one or more file blocks. All file blocks are stored in DataNodes.

Block Replica

A replica is a block copy stored in HDFS. A file block stores multiple replicas for system availability and fault tolerance.

NodeManager

NodeManager executes applications, monitors the usage of resources (including CPUs, memory, disks, and network resources) of applications, and reports the resource usage to the ResourceManager.

ResourceManager

ResourceManager schedules resources required by applications. It provides a scheduling plug-in for allocating cluster resources to multiple queues and

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

applications. The scheduling plug-in schedules resources based on existing capabilities or using the fair scheduling model.

Kafka Partitions

Each topic can be divided into multiple partitions. Each partition corresponds to an appendant log file whose sequence is fixed.

Follower

A follower processes read requests and works with a leader to process write requests. It can also be used as a leader backup. When the leader is faulty, a follower is elected to take over the leader's workload to prevent a single point of failure.

Observer

Observers do not take part in voting for election and write requests. They only process read requests and forward write requests to the leader, improving processing efficiency.

DStream

DStream is an abstract concept provided by Spark Streaming. It is a continuous data stream which is obtained from the data source or the transformed input stream. In essence, a DStream is a series of continuous resilient distributed datasets (RDDs).

Heap Memory

A heap indicates the data area where the Java Virtual Machine (JVM) is running and from which memory for all class instances and arrays is committed. The initial heap memory is controlled by the JVM startup parameter **-Xms**.

- Maximum heap memory: Heap memory that can be committed to a program at most by the system, which is specified by the **-Xmx** parameter.
- Committed heap memory: total heap memory committed by the system for running a program. It ranges from the initial heap memory and the maximum heap memory.
- Used heap memory: heap memory that has been used by a program. It is smaller than the committed heap memory.
- Non-heap memory: memory excluded from the JVM heaps and the memory area for running the JVM. Non-heap memory has the following three memory pools:
  - Code Cache: stores JIT compiled code. Its value is set through the JVM startup parameter **-XX:InitialCodeCacheSize -XX:ReservedCodeCacheSize**. The default value is 240 MB.
  - Compressed Class Space: stores metadata of a pointer. Its value is set through the JVM startup parameter **-XX:CompressedClassSpaceSize**. The default value is 1024 MB.
  - Metaspace: stores metadata. Its value is set through the JVM startup parameter **-XX:MetaspaceSize -XX:MaxMetaspaceSize**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Maximum non heap memory: non-heap memory committed to a program at most by the system. The value is the sum of the code buffer, space of compressed class pointers, and maximum metaspace.
- Committed non-heap memory: total non-heap memory committed by the system for running a program. It ranges from the initial non-heap memory and the maximum non-heap memory.
- Used non heap memory: non heap memory that has been used by a program. It is smaller than the committed non heap memory.

Hadoop

Hadoop is a distributed system framework. It allows users to develop distributed applications using high-speed computing and storage provided by clusters without knowing the underlying details of the distributed system. It can also reliably and efficiently process massive amounts of data in scalable, distributed mode. Hadoop is reliable because it maintains multiple work data duplicates, enabling distributed processing for failed nodes. Hadoop is highly efficient because it processes data in parallel mode. Hadoop is scalable because it can process petabytes of data. Hadoop consists of HDFS, MapReduce, HBase, and Hive.

Role

A role is an element of a service. A service contains one or multiple roles. Services are installed on servers through roles so that they can run properly.

Cluster

A cluster is computer technology that enables multiple servers to work as one server. Clusters improve the stability, reliability, and data processing or service capability of the system. For example, clusters can prevent single point of failures (SPOFs), share storage resources, reduce system load, and improve system performance.

Instance

An instance is formed when a service role is installed on the host. A service has one or more role instances.

Metadata

Metadata is data that provides information about other data and is also called media data or relay data. It is used to define data properties, specify data storage locations, track historical data, retrieve resources, and record files.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2 Preparations

2.1 Creating an IAM User and Granting MRS Permissions

Use IAM to implement fine-grained permission control over your MRS. With IAM, you can:

- Create IAM users under your cloud account for employees based on your enterprise's organizational structure so that each employee is allowed to access MRS resources using their unique security credential (IAM user).
- Grant only the permissions required for users to perform a specific task.
- Entrust a cloud account or cloud service to perform efficient O&M on your MRS resources.

If your cloud account does not require IAM users, skip this section.

This section describes the procedure for granting permissions (see **Figure 2-1**).

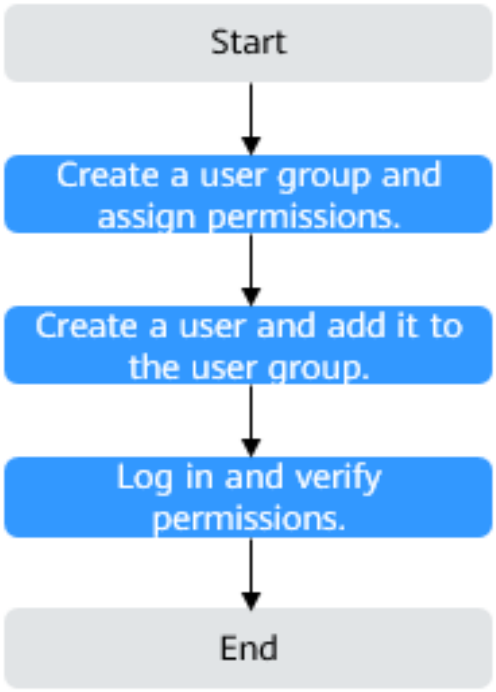
Prerequisites

Before granting permissions to a user group, please familiarize yourself with the MRS permissions that can be added to the user group and select them according to your actual needs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Process Flow

Figure 2-1 Process for granting MRS permissions



1. Create a User Group and Assigning Permissions.

   Create a user group on the IAM console, and assign MRS permissions to the group.

2. Create a user and add it to a user group.

   Create a user on the IAM console and add the user to the group created in 1.

3. Log in and verify permissions.

   Log in to the console by using the user created, and verify that the user has the granted permissions.

   - Choose **Service List** > **Analytics** > **MapReduce Service**. Click **Create Cluster** on the MRS console. If you fail to create an MRS cluster (assume that you only have the **MRS ReadOnlyAccess** permission), the **MRS ReadOnlyAccess** policy has taken effect.
   - Choose any other service in **Service List**. If a message appears indicating that you have insufficient permissions to access the service, the **MRS ReadOnlyAccess** policy has already taken effect.

IAM permissions of MRS

By default, new IAM users lack permissions assigned. Add a user to one or more groups, and attach permissions policies or roles to these groups. Users inherit permissions from the groups to which they are added and can perform specified operations on cloud services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS is a project-level service deployed for specific regions. To assign permissions to a user group, specify **Scope** as **Region-specific projects** and select projects in the corresponding region for the permissions to take effect. If **All projects** is selected, the permissions will take effect for the user group in all region-specific projects. When accessing MRS, the users need to switch to the authorized region.

You can grant permissions by using roles and policies.

- Roles: A coarse-grained IAM authorization strategy to assign permissions based on user responsibilities. Only a limited number of service-level roles are available. Certain roles rely on other roles to take effect. When assigning them, ensure all dependencies are granted. Roles are not an ideal choice for fine-grained authorization and secure access control.
- Policies: A fine-grained authorization mechanism that defines permissions required to perform operations on specific cloud resources under certain conditions. This type of authorization is more flexible and is ideal for secure access control. For example, you can grant MRS users only the permissions for performing specified operations on MRS clusters, such as creating a cluster and querying a cluster list rather than deleting a cluster. Most policies define permissions based on APIs.

Table 2-1 lists all the default system policies supported by MRS.

Table 2-1 MRS system policies

Policy	Description	Type
MRS FullAccess	Administrator permissions for MRS. Users with these permissions can perform all operations on MRS resources.	Fine-grained policy
MRS CommonOperations	Common user permissions for MRS. Users with these permissions can use MRS but cannot add or delete resources.	Fine-grained policy
MRS ReadOnlyAccess	Read-only permission for MRS. Users with these permissions can only view MRS resources.	Fine-grained policy
MRS Administrator	Operation permissions: • All operations on MRS • Permissions of the **Tenant Guest** and **Server Administrator** policies, which must also be granted to the users	RBAC policy

Table 2-2 lists the common operations supported by each system-defined policy of MRS. Select the required policy as needed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 2-2 Common operations supported by each system-defined policy

Operation	MRS FullAccess	MRS CommonOperations	MRS ReadOnlyAccess	MRS Administrator
Creating a cluster	√	x	x	√
Resizing a cluster	√	x	x	√
Upgrading node specifications	√	x	x	√
Deleting a cluster	√	x	x	√
Querying cluster details	√	√	√	√
Listing all clusters	√	√	√	√
Configuring an auto scaling rule	√	x	x	√
Querying a host list	√	√	√	√
Querying operation logs	√	√	√	√
Creating and executing a job	√	√	x	√
Stopping a job	√	√	x	√
Deleting a single job	√	√	x	√
Deleting jobs in batches	√	√	x	√
Querying job details	√	√	√	√

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	MRS FullAccess	MRS CommonOperations	MRS ReadOnlyAccess	MRS Administrator
Querying a job list	√	√	√	√
Creating a folder	√	√	x	√
Deleting a file	√	√	x	√
Querying a file list	√	√	√	√
Operating cluster tags in batches	√	√	x	√
Creating a single cluster tag	√	√	x	√
Deleting a single cluster tag	√	√	x	√
Querying a resource list by tag	√	√	√	√
Querying cluster tags	√	√	√	√
Accessing Manager	√	√	x	√
Querying a patch list	√	√	√	√
Installing a patch	√	√	x	√
Uninstalling a patch	√	√	x	√
Authorizing O&M channels	√	√	x	√
Sharing O&M channel logs	√	√	x	√

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	MRS FullAccess	MRS CommonOperations	MRS ReadOnlyAccess	MRS Administrator
Querying an alarm list	√	√	√	√
Subscribing to alarm notifications	√	√	x	√
Submitting an SQL statement	√	√	x	√
Querying SQL query results	√	√	x	√
Canceling a SQL execution task	√	√	x	√

2.2 Creating a Custom Policy for MRS

Custom policies can be created to supplement the system-defined policies of MRS. For the actions that can be added to custom policies, see **Permissions Policies and Supported Actions** > **Introduction** in MapReduce Service API Reference.

You can create custom policies in either of the following ways:

- Visual editor: Select cloud services, actions, resources, and request conditions. This does not require knowledge of policy syntax.

  Parameters for creating a custom policy for MRS:

  - **Select service**: Select **MapReduce Service (MRS)**.
  - **Select action**: Set it as required.
  - **(Optional) Select resource**: Select **Specific** for **Resources**. If you select **Specify resource path** for **cluster**, **Path** is the ID of the specified MRS cluster. For details, see **MRS Resource Objects**.
- JSON: Create a JSON policy or edit an existing one.

NOTE

Custom policy modifications do not take effect immediately. You need to wait about 15 minutes.

The following section contains examples of common MRS custom policies.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Example Custom Policies

- Example 1: Allowing users to create MRS clusters only

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
            "mrs:cluster:create",
            "ecs:*:*",
            "bms:*:*",
            "evs:*:*",
            "vpc:*:*",
            "smn:*:*"
         ]
      }
   ]
}
```

- Example 2: Allowing users to modify MRS clusters.

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
            "mrs:cluster:resize"
         ]
      }
   ]
}
```

- Example 3: Allowing users to create a cluster, create and execute a job, and delete a single job, but denying cluster deletion

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
            "mrs:cluster:create",
            "mrs:job:submit",
            "mrs:job:delete"
         ]
      },
      {
         "Effect": "Deny",
         "Action": [
            "mrs:cluster:delete"
         ]
      }
   ]
}
```

- Example 4: Granting users the minimum permission to create an MRS cluster with ECS specifications
  - Add the **ecs:serverKeypairs:get** and **ecs:serverKeypairs:list** permissions to use a key pair during cluster creation.
  - Add the **kms:cmk:list** permission to encrypt data disks during cluster creation.
  - Add the **mrs:alarm:subscribe** permission to enable the alarm function during cluster creation.
  - Add the **rds:instance:list** permission to use external data sources during cluster creation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
{
    "Version": "1.1",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "mrs:cluster:create"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ecs:cloudServers:updateMetadata",
                "ecs:cloudServerFlavors:get",
                "ecs:cloudServerQuotas:get",
                "ecs:servers:list",
                "ecs:servers:get",
                "ecs:cloudServers:delete",
                "ecs:cloudServers:list",
                "ecs:serverInterfaces:get",
                "ecs:serverGroups:manage",
                "ecs:servers:setMetadata",
                "ecs:cloudServers:get",
                "ecs:cloudServers:create"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "vpc:securityGroups:create",
                "vpc:securityGroupRules:delete",
                "vpc:vpcs:create",
                "vpc:ports:create",
                "vpc:securityGroups:get",
                "vpc:subnets:create",
                "vpc:privateIps:delete",
                "vpc:quotas:list",
                "vpc:networks:get",
                "vpc:publicIps:list",
                "vpc:securityGroups:delete",
                "vpc:securityGroupRules:create",
                "vpc:privateIps:create",
                "vpc:ports:get",
                "vpc:ports:delete",
                "vpc:publicIps:update",
                "vpc:subnets:get",
                "vpc:publicIps:get",
                "vpc:ports:update",
                "vpc:vpcs:list"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "evs:quotas:get",
                "evs:types:get"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "bms:serverFlavors:get"
            ]
        }
    ]
}
```

- Example 5: Granting users the minimum permission to create an MRS cluster with BMS specifications

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Add the **ecs:serverKeypairs:get** and **ecs:serverKeypairs:list** permissions to use a key pair during cluster creation.
- Add the **kms:cmk:list** permission to encrypt data disks during cluster creation.
- Add the **mrs:alarm:subscribe** permission to enable the alarm function during cluster creation.
- Add the **rds:instance:list** permission to use external data sources during cluster creation.

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "mrs:cluster:create"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ecs:servers:list",
        "ecs:servers:get",
        "ecs:cloudServers:delete",
        "ecs:serverInterfaces:get",
        "ecs:serverGroups:manage",
        "ecs:servers:setMetadata",
        "ecs:cloudServers:create",
        "ecs:cloudServerFlavors:get",
        "ecs:cloudServerQuotas:get"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "vpc:securityGroups:create",
        "vpc:securityGroupRules:delete",
        "vpc:vpcs:create",
        "vpc:ports:create",
        "vpc:securityGroups:get",
        "vpc:subnets:create",
        "vpc:privateIps:delete",
        "vpc:quotas:list",
        "vpc:networks:get",
        "vpc:publicIps:list",
        "vpc:securityGroups:delete",
        "vpc:securityGroupRules:create",
        "vpc:privateIps:create",
        "vpc:ports:get",
        "vpc:ports:delete",
        "vpc:publicIps:update",
        "vpc:subnets:get",
        "vpc:publicIps:get",
        "vpc:ports:update",
        "vpc:vpcs:list"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "evs:quotas:get",
        "evs:types:get"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
            "bms:servers:get",
            "bms:servers:list",
            "bms:serverQuotas:get",
            "bms:servers:updateMetadata",
            "bms:serverFlavors:get"
        ]
    }
  ]
}
```

- Example 6: Allowing users to create a hybrid ECS and BMS cluster with the minimum permission
  - Add the **ecs:serverKeypairs:get** and **ecs:serverKeypairs:list** permissions to use a key pair during cluster creation.
  - Add the **kms:cmk:list** permission to encrypt data disks during cluster creation.
  - Add the **mrs:alarm:subscribe** permission to enable the alarm function during cluster creation.
  - Add the **rds:instance:list** permission to use external data sources during cluster creation.

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "mrs:cluster:create"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ecs:cloudServers:updateMetadata",
        "ecs:cloudServerFlavors:get",
        "ecs:cloudServerQuotas:get",
        "ecs:servers:list",
        "ecs:servers:get",
        "ecs:cloudServers:delete",
        "ecs:cloudServers:list",
        "ecs:serverInterfaces:get",
        "ecs:serverGroups:manage",
        "ecs:servers:setMetadata",
        "ecs:cloudServers:get",
        "ecs:cloudServers:create"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "vpc:securityGroups:create",
        "vpc:securityGroupRules:delete",
        "vpc:vpcs:create",
        "vpc:ports:create",
        "vpc:securityGroups:get",
        "vpc:subnets:create",
        "vpc:privateIps:delete",
        "vpc:quotas:list",
        "vpc:networks:get",
        "vpc:publicIps:list",
        "vpc:securityGroups:delete",
        "vpc:securityGroupRules:create",
        "vpc:privateIps:create",
        "vpc:ports:get",
        "vpc:ports:delete",
        "vpc:publicIps:update",
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
            "vpc:subnets:get",
            "vpc:publicIps:get",
            "vpc:ports:update",
            "vpc:vpcs:list"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "evs:quotas:get",
            "evs:types:get"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "bms:servers:get",
            "bms:servers:list",
            "bms:serverQuotas:get",
            "bms:servers:updateMetadata",
            "bms:serverFlavors:get"
        ]
    }
  ]
}
```

MRS Resource Objects

Resources are objects in a service. Resources in MRS include clusters. You can select a specific resource by specifying the resource path when creating a custom policy.

Table 2-3 MRS resources and their paths

Resource	Name	Path
cluster	Cluster	[Format] MRS:*:*:cluster:*Cluster ID* [Description] Grant permissions on a specific resource to an IAM user. For example, to grant permissions on a specified MRS cluster to an IAM user, set **cluster** to **Specify resource path** and add the resource path. The cluster ID is used to specify the cluster. To view the cluster ID, perform the following steps: 1. On the MRS console, choose **Active Clusters** in the navigation pane on the left and click the name of the target cluster to access its details page. 2. Click the **Dashboard** tab and obtain the cluster ID in the **Basic Information** area.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3 Quick Start

3.1 Creating and Using a Hadoop Cluster for Offline Analysis

Scenario

This topic describes how to create a Hadoop cluster for offline analysis and how to submit a wordcount job through the cluster client. A wordcount job is a classic Hadoop job that counts words in massive amounts of text.

The Hadoop cluster uses the open-source Hadoop ecosystem components, including YARN for cluster resource management, and Hive and Spark for offline large-scale distributed data storage and compute to provide massive data analysis and query capabilities.

Procedure

Before you start, complete operations described in **Preparations**. Then, follow these steps:

1. **Creating an MRS cluster**: Create a Hadoop analysis cluster of MRS 3.1.5.
2. **Installing the Cluster Client**: Download and install the MRS cluster client.
3. **Preparing Applications and Data**: Prepare the data files required for running the wordcount sample program on the MRS cluster client.
4. **Submitting a Job and Viewing the Result**: Submit a wordcount data analysis job on the cluster client and view the execution result.

Preparations

- You have prepared an IAM user who has the permission to create MRS clusters.

Step 1: Creating an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Search for MapReduce Service in the service list and enter the MRS console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click Create Cluster. The **Quick Config** tab is displayed.

Step 4 Configure the cluster as you need.

Table 3-1 MRS cluster parameters

Parameter	Example Value	Description
Region	-	Region where the MRS resources to be requested belong. MRS clusters in different regions cannot communicate with each other over an intranet. For lower network latency and quick resource access, select the nearest region.
Cluster Name	mrs_demo	Name of the MRS cluster you want to create.
Cluster Type	Custom	A range of clusters that accommodate diverse big data demands. You can select a Custom cluster to run a wide range of analytics components supported by MRS.
Version Type	Normal	Service type of the MRS
Cluster Version	MRS 3.1.5	Version of the MRS cluster. Supported open-source components and their functions vary depending on the cluster version. You are advised to select the latest version.
Component	Hadoop Analysis Cluster	Cluster templates containing preset opensource components you will need for your business.
AZ	AZ 1	Available AZ associated with the cluster region.
VPC	vpc-default	VPC where you want to create the cluster. You can click **View VPC** to view the name and ID. If no VPC is available, create one.
Subnet	subnet-default	Subnet where your cluster belongs. You can access the VPC management console to view the names and IDs of existing subnets in the VPC. If no subnet is created under the VPC, click **Create Subnet** to create one.
Cluster Node	Default value	Cluster node details.
Kerberos Authentication	Disabled	Whether Kerberos authentication is enabled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Username	-	Username for logging in to the cluster management page and the ECS node.
Password	-	User password for logging in to the cluster management page and the ECS node.
Confirm Password	-	Enter the user password again.
Enterprise Project	default	Enterprise project to which the cluster belongs.
Secure Communications	Selected	Select the check box to agree to use the access control rules.

Step 5 Click Create Now. A page is displayed showing that the task has been submitted.

Step 6 Click **Back to Cluster List**. You can view the status of the newly created cluster on the **Active Clusters** page.

Wait for the cluster creation to complete. The initial status of the cluster is **Starting**. After the cluster is created, the cluster status becomes **Running**.

----End

Step 2: Installing the Cluster Client

You need to install a cluster client to connect to component services in the cluster, remotely access the client shell, and submit jobs.

The client can be installed on a node in or outside the cluster. This guide describes how to install the client on the Master1 node in the cluster.

Step 1 Click the MRS cluster name in the cluster list to go to the dashboard page.

Step 2 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

For the first access, click **Manage EIPs** to create an EIP on the EIP console. Go back to the **Access MRS Manager** dialog box, refresh the EIP list, and select the EIP.

Step 3 Select the confirmation check box and click **OK** to log in to the FusionInsight Manager of the cluster.

The username for logging in to FusionInsight Manager is **admin**, and the password is the one configured during cluster creation.

Step 4 On the displayed **Homepage** page, click ··· next to the cluster name and click **Download Client** to download the cluster client.

In the **Download Cluster Client** dialog box, set the following parameters:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Set **Select Client Type** to **Complete Client**.
- For **Select Platform Type**, select the architecture of the node where the client is to be installed, for example, **x86_64**.

  To check the architecture of a node in the cluster, click **Hosts** on FusionInsight Manager navigation pane on the top and click the target node name to go to the basic information page.
- Retain the default path for **Save to Path**. The generated file will be saved in the **/tmp/FusionInsight-Client** directory on the active OMS node (usually the **Master1** node) of the cluster.

Click **OK** and wait until the client software is generated.

Step 5 Go back to the MRS console and click the cluster name in the cluster list. Go to the **Nodes** tab, click the name of the node that contains **master1**. In the upper right corner of the ECS details page, click **Remote Login** to log in to the **Master1** node.

Step 6 Log in to the **Master1** node as user **root**. The password is the one you set for the **root** user during cluster creation.

Step 7 Switch to the directory where the client software package is stored and decompress the package.

cd /tmp/FusionInsight-Client/

tar -xvf FusionInsight_Cluster_1_Services_Client.tar

tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar

Step 8 Go to the directory where the installation package is stored and install the client.

cd FusionInsight_Cluster_1_Services_ClientConfig

Install the client to a specified directory. (If the directory exists, it must be empty.)

For example, if the client is installed in the **/opt/client** directory, run the following command:

./install.sh /opt/client

Wait until the client installation is complete.

```
...
... component client is installed successfully
...
```

----End

Step 3: Preparing Applications and Data

You can run the wordcount sample program preset in the cluster client on the created cluster, or develop a big data application and upload it to the cluster.

This topic uses the wordcount sample program on the MRS cluster client as an example. You need to prepare the data files required for running the wordcount sample program.

Step 1 Log in to the **Master1** node as user **root**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Prepare data files.

For example, the file names are **wordcount1.txt** and **wordcount2.txt**, and the content is as follows:

vi /opt/wordcount1.txt

```
hello word
hello wordcount
```

vi /opt/wordcount2.txt

```
hello mapreduce
hello hadoop
```

Step 3 Switch to the client installation directory, configure environment variables, and create an HDFS directory for storing sample data, for example, **/user/example/input**.

cd /opt/client

source bigdata_env

hdfs dfs -mkdir -p /user/example/input

Step 4 Upload the sample data to HDFS.

hdfs dfs -put /opt/wordcount1.txt /user/example/input

hdfs dfs -put /opt/wordcount2.txt /user/example/input

----End

Step 4: Submitting a Job and Viewing the Result

Step 1 Log in to the client node (**Master1**) as user **root**.

Step 2 Submit a wordcount job, read source data for analysis, and output the execution result to the HDFS.

cd /opt/client

source bigdata_env

hadoop jar HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.1-*.jar wordcount "/user/example/input/*" "/user/example/output/"

```
...
    File Input Format Counters
        Bytes Read=56
    File Output Format Counters
        Bytes Written=48
```

NOTE

- **/user/example/output/** indicates the address for storing job output files on the HDFS. Set it to a directory that does not exist.
- The name of the **hadoop-mapreduce-examples-3.3.1-*.jar** file varies depending on the cluster client version. Use the actual name.

Step 3 Query job execution results.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Run the following command to view the job output file:

   hdfs dfs -ls /user/example/output/

   ```
   ...
   ... /user/example/output/_SUCCESS
   ... /user/example/output/part-r-0000
   ```

2. The output is saved in the HDFS file system. You can run a command to download the output to the local PC and view it.

   The following command is an example:

   hdfs dfs -get /user/example/output/part-r-00000 /opt

   cat /opt/part-r-00000

   The content of the **part-r-00000** file is as follows:

   ```
   hadoop    1
   hello    4
   mapreduce    1
   word    1
   wordcount    1
   ```

Step 4 View job run logs.

1. Log in to FusionInsight Manager of the target cluster as user **admin** and choose **Cluster** > **Services** > **Yarn**.
2. Click the **ResourceManager(xxx,Active)** link in the row where the **ResourceManager Web UI** is.
3. On the **All Applications** page, click the ID of the target job to view the job details.

   NOTE

   On the **All Applications** page, you can confirm a task based on the task submission time and the user name that submits the task.

----End

3.2 Creating and Using a Kafka Cluster for Stream Processing

Scenario

This topic helps you create a stream analysis cluster from scratch and generate and consume messages in a Kafka topic.

A Kafka cluster provides a message system with high throughput and scalability. It is widely used for log collection and monitoring data aggregation. Kafka is efficient in streaming data ingestion and real-time data processing and storage.

Procedure

Before you start, complete operations described in **Preparations**. Then, follow these steps:

1. **Creating an MRS Cluster**: Create a real-time analysis cluster of MRS 3.2.0-LTS.1.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. **Installing the Cluster Client**: Download and install the MRS cluster client.
3. **Using the Kafka Client to Create a Topic**: Create a topic on the Kafka client.
4. **Managing Messages in a Kafka Topic**: Consume messages in a created topic on the Kafka client.

Preparations

- You have prepared an IAM user who has the permission to create MRS clusters.

Step 1: Creating an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Search for MapReduce Service in the service list and enter the MRS console.

Step 3 Click Create Cluster. The **Quick Config** tab is displayed.

Step 4 Configure the cluster as you need.

Table 3-2 MRS cluster parameters

Parameter	Example Value	Description
Region	-	Region where the MRS resources to be requested belong. MRS clusters in different regions cannot communicate with each other over an intranet. For lower network latency and quick resource access, select the nearest region.
Cluster Name	**mrs_demo**	Name of the MRS cluster you want to create.
Cluster Type	Custom	A range of clusters that accommodate diverse big data demands. You can select a Custom cluster to run a wide range of analytics components supported by MRS.
Version Type	LTS	Version of the MRS cluster. Supported open-source components and their functions vary depending on the cluster version. You are advised to select the latest version.
Cluster Version	**MRS 3.2.0-LTS.1**	Service type of the MRS
Component	Real-time Analysis Cluster	Cluster templates containing preset opensource components you will need for your business.
AZ	AZ 1	Available AZ associated with the cluster region.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
VPC	**vpc-default**	VPC where you want to create the cluster. You can click **View VPC** to view the name and ID. If no VPC is available, create one.
Subnet	**subnet-default**	Subnet where your cluster belongs. You can access the VPC management console to view the names and IDs of existing subnets in the VPC. If no subnet is created under the VPC, click **Create Subnet** to create one.
Cluster Node	Default value	Cluster node details.
Kerberos Authentication	Disabled	Whether Kerberos authentication is enabled.
Username	-	Username for logging in to the cluster management page and the ECS node.
Password	-	User password for logging in to the cluster management page and the ECS node.
Confirm Password	-	Enter the user password again.
Enterprise Project	**default**	Enterprise project to which the cluster belongs.
Secure Communications	Selected	Select the check box to agree to use the access control rules.

Step 5 Click Create Now. A page is displayed showing that the task has been submitted.

Step 6 Click **Back to Cluster List**. You can view the status of the newly created cluster on the **Active Clusters** page.

Wait for the cluster creation to complete. The initial status of the cluster is **Starting**. After the cluster is created, the cluster status becomes **Running**.

----End

Step 2: Installing the Cluster Client

You need to install a cluster client to connect to component services in the cluster and submit jobs.

You can install the client on a node in or outside the cluster. This topic installs the client on the **Master1** node as an example.

Step 1 Click the MRS cluster name in the cluster list to go to the dashboard page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

For the first access, click **Manage EIPs** to create an EIP on the EIP console. Go back to the **Access MRS Manager** dialog box, refresh the EIP list, and select the EIP.

Step 3 Select the confirmation check box and click **OK** to log in to the FusionInsight Manager of the cluster.

The username for logging in to FusionInsight Manager is **admin**, and the password is the one configured during cluster creation.

Step 4 On the displayed **Homepage** page, click ··· next to the cluster name and click **Download Client** to download the cluster client.

In the **Download Cluster Client** dialog box, set the following parameters:

- Set **Select Client Type** to **Complete Client**.
- Retain the default value for **Platform Type**, for example, **x86_64**.
- Retain the default path for **Save to Path**. The generated file will be saved in the **/tmp/FusionInsight-Client** directory on the active OMS node of the cluster.

Click **OK** and wait until the client software is generated.

Step 5 Go back to the MRS console and click the cluster name in the cluster list. Go to the **Nodes** tab, click the name of the node that contains **master1**. In the upper right corner of the ECS details page, click **Remote Login** to log in to the **Master1** node.

Step 6 Log in to the **Master1** node as user **root**. The password is the one you set for the **root** user during cluster creation.

Step 7 Switch to the directory where the client software package is stored and decompress the package.

cd /tmp/FusionInsight-Client/

tar -xvf FusionInsight_Cluster_1_Services_Client.tar

tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar

Step 8 Go to the directory where the installation package is stored and install the client.

cd FusionInsight_Cluster_1_Services_ClientConfig

Install the client to a specified directory (an absolute path), for example, **/opt/client**.

./install.sh /opt/client

```
...
... component client is installed successfully
...
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

A client installation directory will be automatically created if it does not exist. If there is such directory, it must be empty. The directory name cannot contain spaces. The client installation directory can contain only uppercase letters, lowercase letters, digits, and underscores (_).

----End

Step 3: Using the Kafka Client to Create a Topic

Step 1 In the cluster list, click the name of the target cluster. The dashboard tab is displayed.

Step 2 On the displayed page, click **Synchronize** next to **IAM User Sync**. In the displayed dialog box, select **All**, and click **Synchronize**. Wait until the synchronization task is complete.

Step 3 Go to the **Components** tab, click **ZooKeeper**, and then click the **Instances** tab. Check and record the IP address of a ZooKeeper quorumpeer role instance.

Step 4 Click **Service Configuration** and check the value of **clientPort**, which indicates the ZooKeeper client connection port.

Step 5 Click **Service ZooKeeper** to return to the component list.

Step 6 Click **Kafka**, and then the **Instances** tab. Check and record the IP addresses of a Kafka Broker instance.

Step 7 Click **Service Configuration** and check the value of **port**, which indicates the port for connecting to Kafka Broker.

Step 8 Log in to the node (**Master1**) where the MRS client is located as user **root**.

Step 9 Switch to the client installation directory and configure environment variables.

cd /opt/client

source bigdata_env

Step 10 Create a Kafka topic.

kafka-topics.sh --create --zookeeper *IP address of ZooKeeper role instance*:*ZooKeeper client connection port* **/kafka --partitions 2 --replication-factor 2 --topic** *Topic name*

The following is an example:

kafka-topics.sh --create --zookeeper 192.168.21.234:2181/kafka --partitions 2 --replication-factor 2 --topic Topic1

If the following information is displayed, the topic is created:

```
Created topic Topic1.
```

----End

Step 4: Managing Messages in the Kafka Topic

Step 1 Log in to the node (**Master1**) where the MRS client is deployed as user **root**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Switch to the client installation directory and configure environment variables.

cd /opt/client

source bigdata_env

Step 3 Generate a message in Topic1.

kafka-console-producer.sh --broker-list *IP address of the node where the Kafka Broker role is deployed*:*Broker connection port* **--topic** *Topic name* **--producer.config /opt/hadoopclient/Kafka/kafka/config/producer.properties**

For the IP address and port number of the node where the Kafka Broker instance is deployed, see **Step 6** and **Step 7** in **Step 3: Using the Kafka Client to Create a Topic**.

The following is an example:

kafka-console-producer.sh --broker-list 192.168.21.21:9092 --topic Topic1 --producer.config /opt/client/Kafka/kafka/config/producer.properties

Step 4 Open a new client connection window.

cd /opt/client

source bigdata_env

Step 5 Consume messages in Topic1.

kafka-console-consumer.sh --topic *Topic name* **--bootstrap-server** *IP address of the node where the Kafka Broker role is deployed*:*Broker connection port* **--consumer.config /opt/client/Kafka/kafka/config/consumer.properties**

The following is an example:

kafka-console-consumer.sh --topic Topic1 --bootstrap-server 192.168.21.21:9092 --consumer.config /opt/client/Kafka/kafka/config/consumer.properties

Step 6 Enter some content in the command line that generates messages in **Step 3**. The content is used as the messages generated by the producer. Press **Enter** to send the message.

The following is an example:

```
>aaa
>bbb
>ccc
```

To stop generating messages, press **Ctrl+C** to exit.

Step 7 In the message consuming window of **Step 5**, check whether the messages are consumed.

```
aaa
bbb
ccc
```

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3.3 Creating and Using an HBase Cluster for Offline Query

Scenario

This section describes how to create an HBase query cluster with Kerberos authentication disabled from scratch as well as how to create and query HBase tables through the cluster client.

An HBase cluster uses Hadoop and HBase components to provide a column-oriented distributed cloud storage system featuring enhanced reliability, great performance, and elastic scalability. It applies to the storage and distributed computing of massive amounts of data. You can use HBase to build a storage system capable of storing TB- or even PB-level data. With HBase, you can filter and analyze data with ease and get responses in milliseconds, rapidly mining data value.

Procedure

Before you start, complete operations described in **Preparations**. Then, follow these steps:

1. **Creating an MRS Cluster**: Create an HBase query cluster with Kerberos authentication disabled in MRS 3.2.0-LTS.1.
2. **Installing the Cluster Client**: Download and install the MRS cluster client.
3. **Creating a Table Using the HBase Client**: Create a table, insert table data, query the data, and delete the table on the HBase client.

Preparations

- You have prepared an IAM user who has the permission to create MRS clusters.

Step 1: Creating an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Search for MapReduce Service in the service list and enter the MRS console.

Step 3 Click Create Cluster. The **Quick Config** tab is displayed.

Step 4 Configure the cluster as you need.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 3-3 MRS cluster parameters

Parameter	Example Value	Description
Region	-	Region where the MRS resources to be requested belong. MRS clusters in different regions cannot communicate with each other over an intranet. Select a region nearest to you to minimize network latency and accelerate resource access.
Cluster Name	**mrs_demo**	Name of the MRS cluster you want to create.
Cluster Type	Custom	A range of clusters that accommodate diverse big data demands. You can select a Custom cluster to run a wide range of analytics components supported by MRS.
Version Type	LTS	Version of the MRS cluster. Supported open-source components and their functions vary depending on the cluster version.
Cluster Version	**MRS 3.2.0-LTS.1**	Version of the MRS cluster. Supported open-source components and their functions vary depending on the cluster version. You are advised to select the latest version.
Component	HBase Query Cluster	Cluster templates with preset open-source components required for your business.
AZ	AZ 1	Available AZ associated with the cluster region.
VPC	**vpc-default**	VPC where you want to create the cluster. You can click **View VPC** to view the name and ID. If no VPC is available, create one.
Subnet	**subnet-default**	Subnet where your cluster belongs. You can access the VPC management console to view the names and IDs of existing subnets in the VPC. If no subnet is created under the VPC, click **Create Subnet** to create one.
Cluster Node	Default value	Cluster node details.
Kerberos Authentication	Disabled	Whether Kerberos authentication is enabled.
Username	-	Username for logging in to the cluster management page and the ECS node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Password	-	User password for logging in to the cluster management page and the ECS node.
Confirm Password	-	Enter the user password again.
Enterprise Project	**default**	Enterprise project to which the cluster belongs.
Secure Communications	Selected	Select the check box to agree to use the access control rules.

Step 5 Click Create Now. A page is displayed showing that the task has been submitted.

Step 6 Click **Back to Cluster List**. You can view the status of the newly created cluster on the **Active Clusters** page.

Wait for the cluster creation to complete. The initial status of the cluster is **Starting**. After the cluster is created, the cluster status becomes **Running**.

----End

Step 2: Installing the Cluster Client

You need to install a cluster client to connect to component services in the cluster and submit jobs.

You can install the client on a node in or outside the cluster. This topic installs the client on the **Master1** node as an example.

Step 1 Click the MRS cluster name in the cluster list to go to the dashboard page.

Step 2 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

For the first access, click **Manage EIPs** to create an EIP on the EIP console. Go back to the **Access MRS Manager** dialog box, refresh the EIP list, and select the EIP.

Step 3 Select the confirmation check box and click **OK** to log in to the FusionInsight Manager of the cluster.

The username for logging in to FusionInsight Manager is **admin**, and the password is the one configured during cluster creation.

Step 4 On the displayed **Homepage** page, click ··· next to the cluster name and click **Download Client** to download the cluster client.

In the **Download Cluster Client** dialog box, set the following parameters:

- Set **Select Client Type** to **Complete Client**.
- Retain the default value for **Platform Type**, for example, **x86_64**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Retain the default path for **Save to Path**. The generated file will be saved in the **/tmp/FusionInsight-Client** directory on the active OMS node of the cluster.

Click **OK** and wait until the client software is generated.

Step 5 Go back to the MRS console and click the cluster name in the cluster list. Go to the **Nodes** tab, click the name of the node that contains **master1**. In the upper right corner of the ECS details page, click **Remote Login** to log in to the **Master1** node.

Step 6 Log in to the **Master1** node as user **root**. The password is the one you set for the **root** user during cluster creation.

Step 7 Switch to the directory where the client software package is stored and decompress the package.

cd /tmp/FusionInsight-Client/

tar -xvf FusionInsight_Cluster_1_Services_Client.tar

tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar

Step 8 Go to the directory where the installation package is stored and install the client.

cd FusionInsight_Cluster_1_Services_ClientConfig

Install the client to a specified directory (an absolute path), for example, **/opt/client**.

./install.sh /opt/client

```
...
... component client is installed successfully
...
```

NOTE

A client installation directory will be automatically created if it does not exist. If there is such directory, it must be empty. The directory name cannot contain spaces. The client installation directory can contain only uppercase letters, lowercase letters, digits, and underscores (_).

----End

Step 3: Creating a Table Using the HBase Client

Step 1 Log in to the node (**Master1**) where the MRS client is deployed as user **root**.

Step 2 Switch to the client installation directory and configure environment variables.

Switch to the client installation directory.

```
cd /opt/client
```

Configure environment variables.

```
source bigdata_env
```

This section uses a cluster with Kerberos authentication disabled as an example. If Kerberos authentication is enabled, you need to run the following command to authenticate the user:

```
kinit Component service user
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Access the HBase shell CLI.

```
hbase shell
```

Step 4 Create the **user_info** table on the HBase client.

1. Create the **user_info** table.

   ```
   create 'user_info',{NAME => 'i'}
   ```

2. Add data to the **user_info** table.

   ```
   put 'user_info','12005000201','i:name','A'
   put 'user_info','12005000201','i:gender','Male'
   put 'user_info','12005000201','i:age','19'
   put 'user_info','12005000201','i:address','City A'
   put 'user_info','12005000201','i:degree','master'
   put 'user_info','12005000201','i:pose','manager'
   ```

Step 5 Query the HBase table.

1. Query usernames and addresses by user ID.

   ```
   scan 'user_info',
   {STARTROW=>'12005000201',STOPROW=>'12005000201',COLUMNS=>['i:name','i:address']}
   ```

   The query result is as follows:

   ```
   ROW                                  COLUMN
   +CELL
    12005000201                         column=i:address, timestamp=2021-10-30T10:21:42.196, value=City
   A
    12005000201                         column=i:name, timestamp=2021-10-30T10:21:18.594,
   value=A
   1 row(s)
   Took 0.0996 seconds
   ```

2. Query information by username.

   ```
   scan 'user_info',{FILTER=>"SingleColumnValueFilter('i','name',=,'binary:A')"}
   ```

   The query result is as follows:

   ```
   ROW                                  COLUMN
   +CELL
    12005000201                         column=i:address, timestamp=2021-10-30T10:21:42.196, value=City
   A
    12005000201                         column=i:age, timestamp=2021-10-30T10:21:30.777,
   value=19
    12005000201                         column=i:degree, timestamp=2021-10-30T10:21:53.284,
   value=master
    12005000201                         column=i:gender, timestamp=2021-10-30T10:21:18.711,
   value=Male
    12005000201                         column=i:name, timestamp=2021-10-30T10:21:18.594,
   value=A
    12005000201                         column=i:pose, timestamp=2021-10-30T10:22:07.152,
   value=manager
   1 row(s)
   Took 0.2158 seconds
   ```

Step 6 Delete the HBase table.

1. Delete user data from the user information table.

   ```
   delete 'user_info','12005000201','i'
   ```

2. Delete the user information table.

   ```
   disable 'user_info'
   drop 'user_info'
   ```

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3.4 Creating and Using a ClickHouse Cluster for Columnar Store

Scenario

This topic helps you create a ClickHouse cluster from scratch and create and query a ClickHouse table through the cluster client.

ClickHouse is an open-source columnar database oriented to online analysis and processing. It is independent of the Hadoop big data system and features ultimate compression rate and fast query performance.

Procedure

Before you start, complete operations described in **Preparations**. Then, follow these steps:

1. **Creating an MRS Cluster**: Create a ClickHouse cluster of MRS 3.2.0-LTS.1.
2. **Installing the Cluster Client**: Download and install the MRS cluster client.
3. **Creating a Table through the ClickHouse Client**: Create a table on the ClickHouse client and insert data into the table.

Preparations

- You have prepared an IAM user who has the permission to create MRS clusters.

Step 1: Creating an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Search for MapReduce Service in the service list and enter the MRS console.

Step 3 Click Create Cluster. The **Quick Config** tab is displayed.

Step 4 Configure the cluster as you need.

Table 3-4 MRS cluster parameters

Parameter	Example Value	Description
Region	-	Region where the MRS resources to be requested belong. MRS clusters in different regions cannot communicate with each other over an intranet. For lower network latency and quick resource access, select the nearest region.
Cluster Name	**mrs_demo**	Name of the MRS cluster you want to create.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Cluster Type	Custom	A range of clusters that accommodate diverse big data demands. You can select a Custom cluster to run a wide range of analytics components supported by MRS.
Version Type	LTS	Service type of the MRS
Cluster Version	**MRS 3.2.0-LTS.1**	Version of the MRS cluster. Supported open-source components and their functions vary depending on the cluster version. You are advised to select the latest version.
Component	ClickHouse Cluster	Cluster templates containing preset opensource components you will need for your business.
AZ	AZ 1	Available AZ associated with the cluster region.
VPC	**vpc-default**	VPC where you want to create the cluster. You can click **View VPC** to view the name and ID. If no VPC is available, create one.
Subnet	**subnet-default**	Subnet where your cluster belongs. You can access the VPC management console to view the names and IDs of existing subnets in the VPC. If no subnet is created under the VPC, click **Create Subnet** to create one.
Cluster Node	Default value	Cluster node details.
Kerberos Authentication	Disabled	Whether Kerberos authentication is enabled.
Username	-	Username for logging in to the cluster management page and the ECS node.
Password	-	User password for logging in to the cluster management page and the ECS node.
Confirm Password	-	Enter the user password again.
Enterprise Project	**default**	Enterprise project to which the cluster belongs.
Secure Communications	Selected	Select the check box to agree to use the access control rules.

Step 5 Click Create Now. A page is displayed showing that the task has been submitted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Click **Back to Cluster List**. You can view the status of the newly created cluster on the **Active Clusters** page.

Wait for the cluster creation to complete. The initial status of the cluster is **Starting**. After the cluster is created, the cluster status becomes **Running**.

----End

Step 2: Installing the Cluster Client

You need to install a cluster client to connect to component services in the cluster and submit jobs.

You can install the client on a node in or outside the cluster. This topic installs the client on the **Master1** node as an example.

Step 1 Click the MRS cluster name in the cluster list to go to the dashboard page.

Step 2 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

For the first access, click **Manage EIPs** to create an EIP on the EIP console. Go back to the **Access MRS Manager** dialog box, refresh the EIP list, and select the EIP.

Step 3 Select the confirmation check box and click **OK** to log in to the FusionInsight Manager of the cluster.

The username for logging in to FusionInsight Manager is **admin**, and the password is the one configured during cluster creation.

Step 4 On the displayed **Homepage** page, click ··· next to the cluster name and click **Download Client** to download the cluster client.

In the **Download Cluster Client** dialog box, set the following parameters:

- Set **Select Client Type** to **Complete Client**.
- Retain the default value for **Platform Type**, for example, **x86_64**.
- Retain the default path for **Save to Path**. The generated file will be saved in the **/tmp/FusionInsight-Client** directory on the active OMS node of the cluster.

Click **OK** and wait until the client software is generated.

Step 5 Go back to the MRS console and click the cluster name in the cluster list. Go to the **Nodes** tab, click the name of the node that contains **master1**. In the upper right corner of the ECS details page, click **Remote Login** to log in to the **Master1** node.

Step 6 Log in to the **Master1** node as user **root**. The password is the one you set for the **root** user during cluster creation.

Step 7 Switch to the directory where the client software package is stored and decompress the package.

cd /tmp/FusionInsight-Client/

tar -xvf FusionInsight_Cluster_1_Services_Client.tar

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar

Step 8 Go to the directory where the installation package is stored and install the client.

cd FusionInsight_Cluster_1_Services_ClientConfig

Install the client to a specified directory (an absolute path), for example, **/opt/client**.

./install.sh /opt/client

```
...
... component client is installed successfully
...
```

NOTE

A client installation directory will be automatically created if it does not exist. If there is such directory, it must be empty. The directory name cannot contain spaces. The client installation directory can contain only uppercase letters, lowercase letters, digits, and underscores (_).

----End

Step 3: Creating a Table Through the ClickHouse Client

Step 1 Log in to the node (**Master1**) where the MRS client is deployed as user **root**.

Step 2 Switch to the client installation directory and configure environment variables.

cd /opt/client

source bigdata_env

Step 3 Run the **clickhouse client** command to connect to the ClickHouse server.

clickhouse client --host *IP address of the ClickHouseServer instance* **--port** *9000* **--user** *Username* **--password**

NOTE

- To obtain the IP address of the ClickHouseServer instance, log in to FusionInsight Manager of the cluster and choose **Cluster** > **Services** > **ClickHouse** > **Instances**.
- Clusters with Kerberos authentication disabled use non-SSL connections by default. The default connection port is 9000. To view the port number, log in to FusionInsight Manager of the cluster, choose **Cluster** > **Services** > **ClickHouse** > **Configurations**, and search for **tcp_port**.
- If the **--user** and **--password** parameters are not specified, the **default** user is used for logging in to the ClickHouse client by default. If you want to specify the username and password, run the **create user** SQL statement on the ClickHouse client to create a ClickHouse user.

Step 4 Create the **test001** database.

```
create database test001 on cluster default_cluster;
```

Step 5 Create the replication table **test010**.

```
CREATE TABLE test001.test010 on cluster default_cluster
(
`EventDate` DateTime,
`CounterID` UInt32,
`UserID` UInt32,
`ver` UInt16
```

```
)
ENGINE = ReplicatedReplacingMergeTree('/clickhouse/tables/{shard}/test3', '{replica}', ver)
ORDER BY (CounterID, EventDate, intHash32(UserID));
```

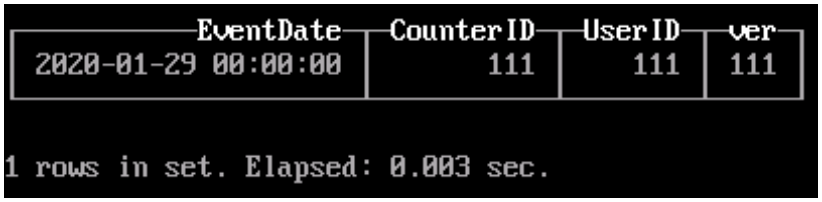
Step 6 Create a replication table **test010** and insert data into the table.

```
insert into test001.test010 values('2020-01-29',111,111,111);
```

Step 7 View data in the **test010** table.

```
select * from test001.test010;
```

Figure 3-1 Viewing data in the **test010** table.



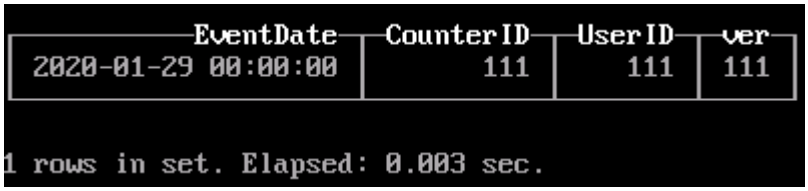
Step 8 Create a distributed table **test001.test010_dir** and insert data into the table.

```
create table test001.test010_dir ON CLUSTER default_cluster as test001.test010 ENGINE =
Distributed(default_cluster, test001, test010, rand());
insert into test001.test010_dir values('2020-01-29',111,111,111);
```

Step 9 Query the distributed table.

```
select * from test001.test010_dir;
```

Figure 3-2 Viewing data in the **test010_dir** table



Step 10 Delete the created database table.

```
drop database test001 on cluster default_cluster no delay;
```

----End

3.5 Creating and Using an MRS Cluster Requiring Security Authentication

Scenario

This topic helps you create a Hadoop analysis cluster that requires Kerberos authentication and submit a wordcount job through the cluster client. A wordcount job is a classic Hadoop job that counts words in massive amounts of text.

The Hadoop cluster uses many open-source Hadoop ecosystem components, including YARN for cluster resource management and Hive and Spark for offline large-scale distributed data storage and compute to provide massive data analysis and query capabilities.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Before you start, complete operations described in **Preparations**. Then, follow these steps:

1. **Creating an MRS Cluster**: Create a Hadoop analysis cluster of MRS 3.2.0-LTS.1 that requires Kerberos authentication.
2. **Creating a Cluster User**: Create a role that has the permission to submit the wordcount job and bind the role to a user on FusionInsight Manager.
3. **Installing the Cluster Client**: Download and install the MRS cluster client.
4. **Preparing Applications and Data**: Prepare the data files required for running the wordcount sample program on the MRS cluster client.
5. **Submitting a Job and Viewing the Result**: Submit a wordcount data analysis job on the cluster client and view the execution result.

Preparations

- You have prepared an IAM user who has the permission to create MRS clusters.

Step 1: Creating an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Search for MapReduce Service in the service list and enter the MRS console.

Step 3 Click Create Cluster. The **Quick Config** tab is displayed.

Step 4 Configure the cluster as you need.

Table 3-5 MRS cluster parameters

Parameter	Example Value	Description
Region	-	Region where the MRS resources to be requested belong. MRS clusters in different regions cannot communicate with each other over an intranet. For lower network latency and quick resource access, select the nearest region.
Cluster Name	**mrs_demo**	Name of the MRS cluster you want to create.
Cluster Type	Custom	A range of clusters that accommodate diverse big data demands. You can select a Custom cluster to run a wide range of analytics components supported by MRS.
Version Type	LTS	Service type of the MRS

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Cluster Version	**MRS 3.2.0-LTS.1**	Version of the MRS cluster. Supported open-source components and their functions vary depending on the cluster version. You are advised to select the latest version.
Component	Hadoop Analysis Cluster	Cluster templates containing preset opensource components you will need for your business.
AZ	AZ 1	Available AZ associated with the cluster region.
VPC	**vpc-default**	VPC where you want to create the cluster. You can click **View VPC** to view the name and ID. If no VPC is available, create one.
Subnet	**subnet-default**	Subnet where your cluster belongs. You can access the VPC management console to view the names and IDs of existing subnets in the VPC. If no subnet is created under the VPC, click **Create Subnet** to create one.
Cluster Node	Default value	Cluster node details.
Kerberos Authentication	Enabled	Whether Kerberos authentication is enabled.
Username	-	Username for logging in to the cluster management page and the ECS node.
Password	-	User password for logging in to the cluster management page and the ECS node.
Confirm Password	-	Enter the user password again.
Enterprise Project	**default**	Enterprise project to which the cluster belongs.
Secure Communications	Selected	Select the check box to agree to use the access control rules.

Step 5 Click Create Now. A page is displayed showing that the task has been submitted.

Step 6 Click **Back to Cluster List**. You can view the status of the newly created cluster on the **Active Clusters** page.

Wait for the cluster creation to complete. The initial status of the cluster is **Starting**. After the cluster is created, the cluster status becomes **Running**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2: Creating a Cluster User

For clusters with Kerberos authentication enabled, perform the following steps to create a user and grant permissions to the user to execute programs.

Step 1 Click the MRS cluster name in the cluster list to go to the dashboard page.

Step 2 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

For the first access, click **Manage EIPs** to create an EIP on the EIP console. Go back to the **Access MRS Manager** dialog box, refresh the EIP list, and select the EIP.

Step 3 Select the confirmation check box and click **OK** to log in to the FusionInsight Manager of the cluster.

The username for logging in to FusionInsight Manager is **admin**, and the password is the one configured during cluster creation.

Step 4 Click **System** in the navigation pane on the top, and click **Permission** > **Role**.

Step 5 Click **Create Role** and set the following parameters.

- Enter a role name, for example, **mrrole**.
- For **Configure Resource Permission**, select the cluster to be operated, choose **Yarn** > **Scheduler Queue** > **root**, and select **Submit** and **Admin** in the **Permission** column. Click the name of the target cluster in the path and then configure other permissions.
- Choose **HDFS** > **File System** > **hdfs://hacluster/**. Locate the row that contains **user**, select **Read**, **Write**, and **Execute** in the **Permission** column, and click **OK**.

Step 6 Click **User** in the navigation pane on the left, and then click **Create** on the displayed page. Set the following parameters.

- Enter a username, for example, **test**.
- Set **User Type** to **Human-Machine**.
- Enter the password in **Password** and enter it again in **Confirm Password**.
- Bind **Manager_viewer** to the **mrrole** role created in **Step 5** to grant permissions.

Step 7 Click **OK**.

----End

Step 3: Installing the Cluster Client

You need to install a cluster client to connect to component services in the cluster and submit jobs.

You can install the clients on a node in or outside the cluster. This topic installs the client on the **Master1** node as an example.

Step 1 Click the MRS cluster name in the cluster list to go to the dashboard page.

Step 2 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For the first access, click **Manage EIPs** to create an EIP on the EIP console. Go back to the **Access MRS Manager** dialog box, refresh the EIP list, and select the EIP.

Step 3 Select the confirmation check box and click **OK** to log in to the FusionInsight Manager of the cluster.

The username for logging in to FusionInsight Manager is **admin**, and the password is the one configured during cluster creation.

Step 4 On the displayed **Homepage** page, click ··· next to the cluster name and click **Download Client** to download the cluster client.

In the **Download Cluster Client** dialog box, set the following parameters:

- Set **Select Client Type** to **Complete Client**.
- Retain the default value for **Platform Type**, for example, **x86_64**.
- Retain the default path for **Save to Path**. The generated file will be saved in the **/tmp/FusionInsight-Client** directory on the active OMS node of the cluster.

Click **OK** and wait until the client software is generated.

Step 5 Go back to the MRS console and click the cluster name in the cluster list. Go to the **Nodes** tab, click the name of the node that contains **master1**. In the upper right corner of the ECS details page, click **Remote Login** to log in to the **Master1** node.

Step 6 Log in to the **Master1** node as user **root**. The password is the one you set for the **root** user during cluster creation.

Step 7 Switch to the directory where the client software package is stored and decompress the package.

cd /tmp/FusionInsight-Client/

tar -xvf FusionInsight_Cluster_1_Services_Client.tar

tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar

Step 8 Go to the directory where the installation package is stored and install the client.

cd FusionInsight_Cluster_1_Services_ClientConfig

Install the client to a specified directory (an absolute path), for example, **/opt/client**.

./install.sh /opt/client

```
...
... component client is installed successfully
...
```

NOTE

A client installation directory will be automatically created if it does not exist. If there is such directory, it must be empty. The directory name cannot contain spaces. The client installation directory can contain only uppercase letters, lowercase letters, digits, and underscores (_).

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4: Preparing Applications and Data

You can run the wordcount sample program preset in the cluster client on the created cluster, or develop a big data application and upload it to the cluster. This topic uses the wordcount sample program preset in the cluster client as an example. You need to prepare the data files required for running the wordcount sample program.

Step 1 Log in to the **Master1** node as user **root**.

Step 2 Prepare data files.

There is no format requirement. For example, the file names are **wordcount1.txt** and **wordcount2.txt**, and the content is as follows:

vi /opt/wordcount1.txt

```
hello word
hello wordcount
```

vi /opt/wordcount2.txt

```
hello mapreduce
hello hadoop
```

Step 3 Switch to the client installation directory, configure environment variables, and create an HDFS directory for storing sample data, for example, **/user/example/input**.

cd /opt/client

source bigdata_env

kinit test (**test** is the username created in **Step 6**. Change the password upon first login.)

hdfs dfs -mkdir -p /user/example

hdfs dfs -mkdir -p /user/example/input

NOTE

The **test** user created in **Step 6** has only the read, write, and execute permissions on the **/user** directory. If the **input** directory is created in a directory other than **/user**, an error message is displayed, indicating that the permission is required. The following is an example:

hdfs dfs -mkdir -p /hbase/input

The following error message is displayed:

```
mkdir: Permission denied: user=test, access=EXECUTE, inode="/hbase":hbase:hadoop:drwxrwx--T
```

Step 4 Upload the sample data to HDFS.

hdfs dfs -put /opt/wordcount1.txt /user/example/input

hdfs dfs -put /opt/wordcount2.txt /user/example/input

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5: Submitting a Job and Viewing the Result

Step 1 Log in to the client node (**Master1**) as user **root**.

Step 2 Submit the wordcount job, read source data for analysis, and output the execution result to the HDFS.

cd /opt/client

source bigdata_env

kinit test

hadoop jar HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.1-*.jar wordcount "/user/example/input/*" "/user/example/output/"

```
...
  File Input Format Counters
    Bytes Read=56
  File Output Format Counters
    Bytes Written=48
```

NOTE

- **/user/example/output/** indicates the address for storing job output files on the HDFS. Set it to a directory that does not exist.
- The name of the **hadoop-mapreduce-examples-3.3.1-*.jar** file varies depending on the cluster client version. Use the actual name.

Step 3 Query job execution results.

1. View the job output file.

   hdfs dfs -ls /user/example/output/

   ```
   ...
   ... /user/example/output/_SUCCESS
   ... /user/example/output/part-r-0000
   ```

2. Save the output in the HDFS file system. You can run a command to download the output to the local PC and view it.

   The following is an example:

   hdfs dfs -get /user/example/output/part-r-00000 /opt

   cat /opt/part-r-00000

   The content of the **part-r-00000** file is as follows:

   ```
   hadoop    1
   hello    4
   mapreduce    1
   word    1
   wordcount    1
   ```

Step 4 View job run logs.

1. Log in to FusionInsight Manager of the cluster as user **test** created in **Step 6** and choose **Cluster** > **Services** > **Yarn**.
2. Click the **ResourceManager(xxx,Active)** link in the row where **ResourceManager Web UI** is.
3. On the **All Applications** page, click the ID of the target job to view the job details.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

On the **All Applications** page, you can confirm a task based on the task submission time and the user name that submits the task.

----End

4 MRS Cluster Planning

4.1 Service Selection

4.1.1 MRS Service Scenarios

MRS consists of multiple big data components, and you can select the cluster type that best fits your service requirements, data types, reliability expectations, and resource budget.

You can **quickly buy** a cluster using the preset cluster template or select the component list and advanced settings to **manually buy** a cluster.

Table 4-1 MRS cluster types

Type	Scenario	Core component
Hadoop analysis cluster	Hadoop cluster uses components in the open source Hadoop ecosystem to analyze and query vast amounts of data. For example, use YARN to manage cluster resources, Hive and Spark to provide offline storage and computing of large-scale distributed data, Spark Streaming and Flink to offer streaming data computing, and Tez to provide a distributed computing framework of directed acyclic graphs (DAGs).	Hadoop, Hive, Spark, Tez, Flink, ZooKeeper, and Ranger

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Scenario	Core component
HBase query cluster	An HBase cluster uses Hadoop and HBase components to provide a column-oriented distributed cloud storage system featuring enhanced reliability, great performance, and elastic scalability. It applies to the storage and distributed computing of massive amounts of data. You can use HBase to build a storage system capable of storing TB- or even PB-level data. With HBase, you can filter and analyze data with ease and get responses in milliseconds, rapidly mining data value.	Hadoop, HBase, ZooKeeper, and Ranger
ClickHouse cluster	ClickHouse is a columnar database management system for online analysis. It features ultimate compression ratio and fast query performance. It is widely used in Internet advertisement, app and web traffic analysis, telecom, finance, and IoT fields.	ClickHouse and ZooKeeper
Real-time analysis cluster	Real-time analysis clusters use Hadoop, Kafka, Flink, and ClickHouse components to provide a system for collection, real-time analysis, and query of data at scale.	Hadoop, Kafka, Flink, ClickHouse, ZooKeeper, and Ranger

Big data is ubiquitous in our lives. Huawei Cloud MRS is suitable for big data processing in the industries such as the Internet of things (IoT), e-commerce, finance, manufacturing, healthcare, and energy.

The following are some typical service scenarios of big data processing:

Offline Data Lake

Offline data processing involves the analysis and processing of large volumes of data to generate results for various data applications.

Offline data processing has low requirements on processing time. However, a large amount of data needs to be processed, which occupies a large number of compute and storage resources. Generally, offline processing is implemented through Hive/SparkSQL or MapReduce/Spark.

- Massive data storage

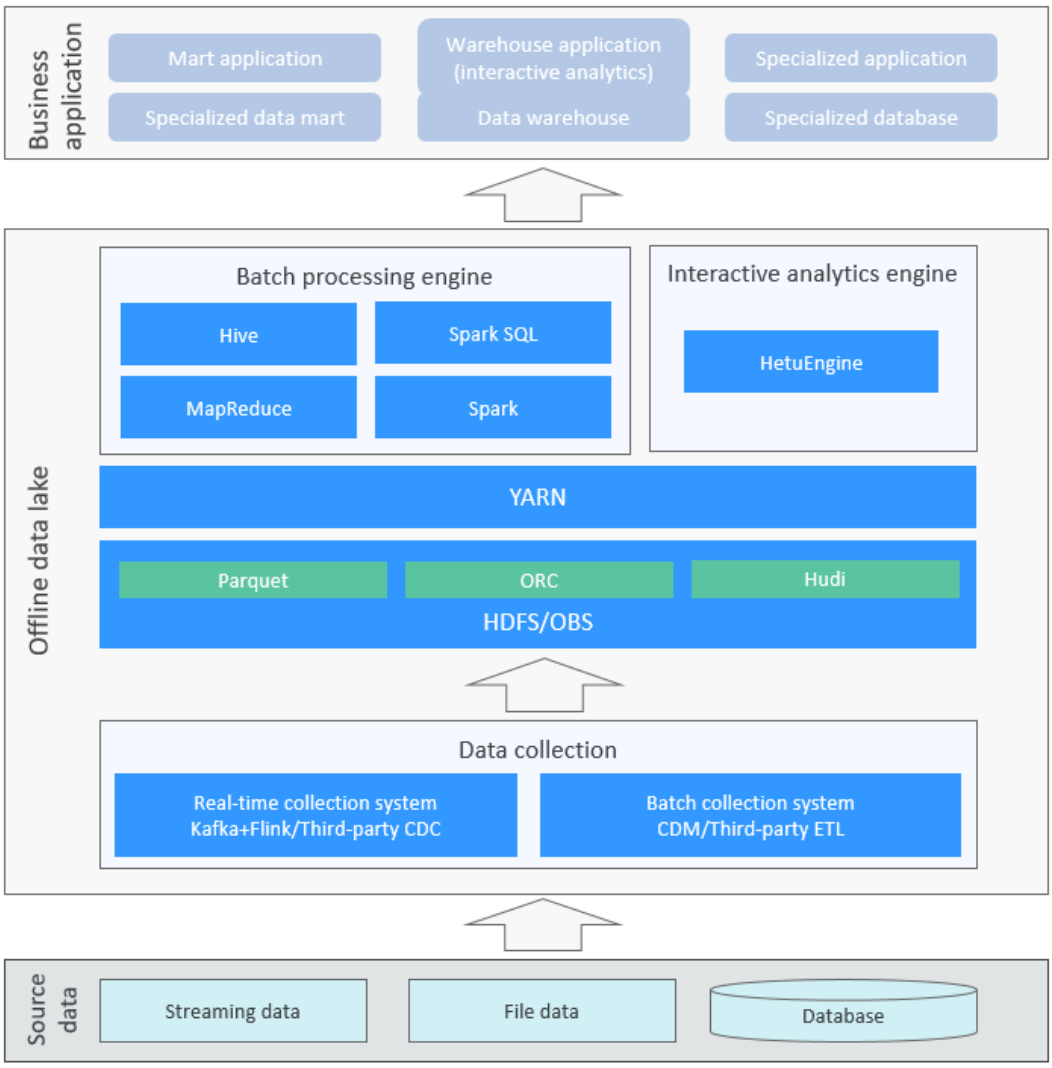
  HDFS supports distributed storage of TB- and PB-level data in clusters. It uses the copy mechanism (three copies by default) to ensure data reliability and is well-suited for storing logs and unstructured data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Batch processing framework

  Complex computing tasks can be divided into two phases: "Map" and "Reduce". This framework supports large-scale parallel computing, such as sorting, aggregation, and text analysis.

Figure 4-1 Offline data analysis



- Source data: includes streaming data, batch file data, and databases.
- Real-time collection system: collects data in real time, for example, using Kafka, Flink, and third-party CDC tools.
- Batch collection system: collects offline data in batches using DataArts Studio-CDM and third-party ETL tools.
- Batch processing engine: implements high-performance execution of offline batch processing jobs.
  - Hive: a conventional SQL batch processing engine used to process SQL-like batch processing jobs. Its performance is stable in the case of large amounts of data, but its processing speed is slow.
  - MapReduce: a conventional batch processing engine used to process non-SQL batch processing jobs, especially batch processing jobs of data

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

mining and machine learning. It is widely used and has stable performance in the case of massive amounts of data, but its processing speed is slow.

- Spark SQL: a new SQL batch processing engine used to process SQL-like batch processing jobs. It is applicable to scenarios with huge amounts of data, and its processing speed is fast.
- Spark: a new batch processing engine used to process non-SQL batch processing jobs, especially batch processing jobs of data mining and machine learning. It is applicable to scenarios with massive amounts of data, and its processing speed is fast.
- YARN: resource scheduling engine that provides resource scheduling capabilities for various batch processing engines. This engine is the basis of multi-tenant resource allocation.
- HDFS/OBS: distributed file system that provides data storage for various batch processing engines. It can store data in various file formats, such as Parquet, ORC, and Hudi.

- Interactive analytics engine: implements interactive queries using HetuEngine. It is suitable for scenarios with massive amounts of data, medium and high concurrency, and multi-tenant shared queries.
- Service application: queries and uses batch processing results. It is customized and developed by upper-layer services.

Real-Time Data Lake

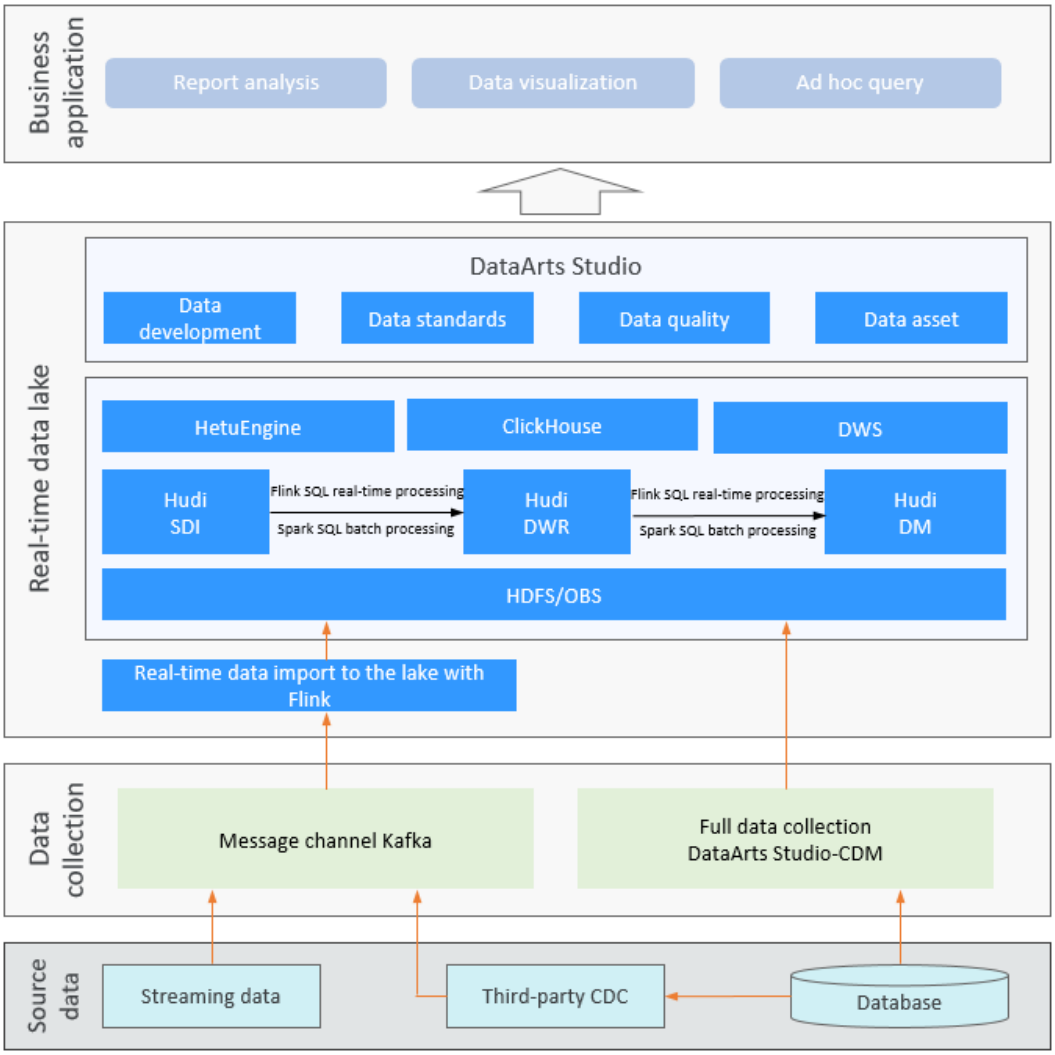
In a real-time data lake, the traditional T+1 data processing mode of offline big data processing (also called offline data lake) is enhanced to achieve minute-level data processing.

The core of a real-time data lake is the data lake. You can implement end-to-end data processing, including real-time data ingestion to a data lake, converged batch and stream data processing within a data lake, and real-time data queries.

- Supports data update and automatic merge.
- Supports streaming processing and reads new data in real time.
- Supports read and write of incremental data. The processing of new data does not require full table scan.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 4-2 Real-time data lake



- Source data: TP databases or streaming message data that can be written to Kafka.
- Data collection: synchronizes data or imports data to the data lake in real time.
  - Kafka: real-time message pipeline used to receive real-time streaming data or changed TP data collected by third parties.
  - DataArts Studio-CDM: migrates all historical data. After all historical data is migrated, CDL is used to synchronize incremental data in real time.
- Real-time data lake: implements high-performance, real-time job processing.
  - Hudi: a data lake storage format that provides the ability to update and delete data and consume new data on top of HDFS.
  - Flink: real-time stream processing engine that can implement converged batch and stream processing based on the real-time incremental query capabilities of Hudi.
  - Spark SQL: batch processing engine that implements Hudi-based batch processing. You can select this engine if real-time service processing is not required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - HDFS/OBS: distributed file system, which provides data storage for various batch processing engines and can store data in various formats.
- Data Lake Insight (DLI): used for efficient access to processed data in the lake.
  - HetuEngine: provides efficient interactive analysis of data within the lake.
  - ClickHouse: MPP database, which is used to store and query data in the mart and provides efficient aggregation analysis for large wide tables.
  - DWS: MPP database, which provides traditional data warehouse capabilities.
- DataArts Studio: a tool set for big data development.

  Provides a UI for visualized data development, data standards design, data quality job design, and data asset display and search, significantly improving the job development efficiency on the data lake.
- Business application: queries and uses data in the real-time data lake. It is customized and developed by upper-layer services.

Real-Time Stream Processing

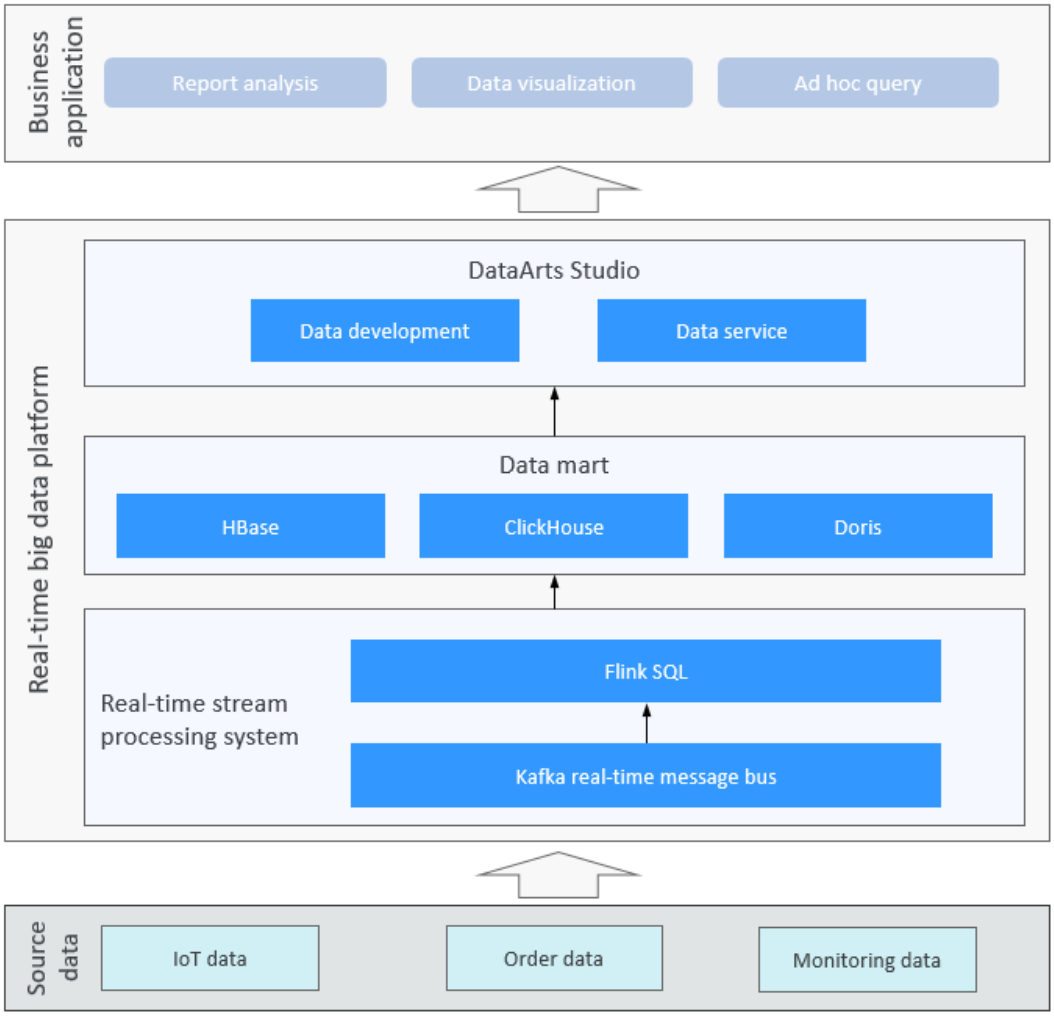
In big data scenarios, real-time stream processing is widely used, just like batch processing. Compared with batch processing, real-time stream processing greatly improves the data processing timeliness, bringing more agile service experience.

A real-time stream processing system must feature low latency, good scalability, and high availability, which can reduce the data processing latency of traditional batch processing services from days or hours to minutes or even seconds.

- Data is collected in real time and sent to the processing system as streams.
- The stream processing engine can continuously process received data in real time.
- The stream processing engine continuously generates calculation results based on received data.
- The system can interconnect with multiple data storage systems and deliver processing results to the data destination.
- The data storage system allows continuous data write and low-latency data query in real time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 4-3 Real-time stream processing



- Source data: streaming message data written to Kafka, which is the data streams that increase continuously.
- Real-time stream processing system: synchronizes data or imports data to the data lake in real time.
  - Kafka: real-time message pipeline, which is used to receive real-time streaming data and serves as the unified data source for real-time stream processing.
  - Flink SQL: real-time stream processing. It reads data from Kafka in real time, performs complex real-time computations, and sends the results to downstream systems, achieving millisecond-level real-time stream processing.
- Data mart: enables efficient access to the data processed through real-time stream processing.
  - HBase: exact query engine, which provides high-concurrency and low-latency access to massive amounts of data. It can quickly retrieve target data from huge amounts of data based on globally unique indexes.
  - ClickHouse: MPP database, which is used to store and query data in the mart and provides efficient aggregation analysis for large wide tables.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Doris: MPP database, which provides a real-time data warehouse.
- DataArts Studio: a tool set for big data development.
  - Data development: provides visualized task development and scheduling capabilities for real-time stream data, and supports real-time processing tasks.
  - Data service: allows you to develop and open APIs for data services. Queries in the data mart are defined as data service APIs to provide API-based data service capabilities for upper-layer applications.
- Business application: queries and uses data in the real-time data mart. It is customized and developed by upper-layer services.

Wide Table Data Mart

In traditional data mart implementations, real-time data and offline data are typically separated, and metrics are often pre-aggregated for custom development, which can lead to inflexibility, low timeliness, and increased complexity.

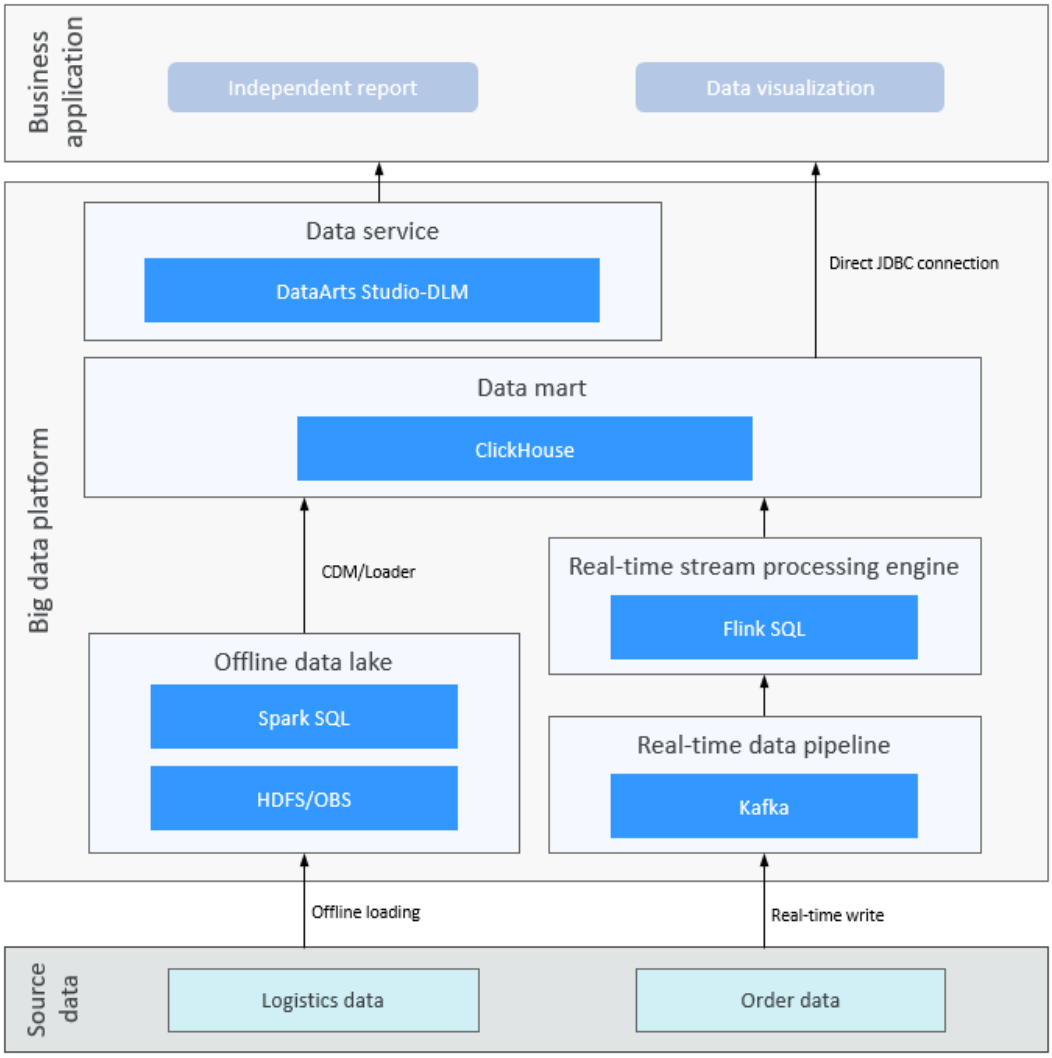
The ClickHouse-based real-time data mart solution normalizes offline and real-time data. Multi-dimensional aggregation queries can be performed within seconds on detailed data without pre-aggregation. This solution features real time, high efficiency, and good flexibility.

ClickHouse is a high-performance, column-oriented database management system (DBMS) for online analytical processing (OLAP). It provides optimal query performance with the ability to process hundreds of millions of rows per second on a single node and distributed scalability.

- Real-time data write, sub-second aggregation query delay, and high timeliness
- No preprocessing required and data query based on detailed wide tables, which is cost-effective
- Multi-dimensional flexible query with different combinations and full self-service metric query, providing high flexibility

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 4-4 Wide table data mart



- Source data: data generated by service systems, which can be offline file data, relational database data, or real-time streaming data.
- Big data platform: core data processing platform for the real-time wide table data mart, which consists of modules such as offline data processing, real-time data processing, and the wide table data mart.
  - Real-time message pipeline: receives real-time streaming data and serves as a unified pipeline for real-time data ingestion.
  - Real-time stream processing engine: reads data from Kafka in real time, performs complex real-time computations, and sends the results to downstream systems, achieving millisecond-level real-time stream processing.
  - Offline data lake: collects and processes offline data from service systems, processes wide tables, uses HDFS or OBS for data storage, and uses Spark SQL as the data processing engine.
  - Wide table data mart: stores and queries mart data, and provides efficient aggregation and analysis of large wide tables. ClickHouse is the core component of the wide table data mart.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Data service: allows you to develop and open APIs for data services. Queries in the data mart are defined as data service APIs to provide API-based data service capabilities for upper-layer applications.
  - Direct JDBC connection: ClickHouse provides standard JDBC APIs. When BI tools are interconnected, you can also use the JDBC APIs to connect to ClickHouse for data query.
- Business application: queries and uses data in the real-time data lake. It is customized and developed by upper-layer services.

4.1.2 MRS Cluster Node Types

An MRS cluster consists of multiple ECSs. The system manages nodes in node groups based on specifications. Nodes in the same node group use same ECS specifications.

Nodes in a cluster can be classified into Master nodes, Core nodes, and Task nodes based on the roles of components deployed on the nodes.

Table 4-2 Cluster node types

Node Type	Functions
Master node	MRS cluster management node. OMS server is deployed on the node to manage and monitor the cluster. After creating the MRS cluster, the node containing **master1** in its name functions as the Master1 node, and the one with **master2** in its name functions as the Master2 node. You can log in to a Master node either using VNC on the ECS management console or using SSH. After logging in to the Master node, you can access other nodes without entering passwords. The system automatically deploys the Master nodes in active/standby mode and supports the high availability (HA) feature for MRS cluster management. If the active management node fails, the standby management node switches to the active state and takes over services. To determine whether the Master1 node is the active management node, see **Checking MRS Active/Standby Management Nodes**.
Core node	Work node of an MRS cluster. It processes and analyzes data and stores process data. In the **Nodes** tab of the cluster details page, the nodes in the node group whose **Node Type** includes **Core** are Core nodes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Node Type	Functions
Task node	Compute node. When the compute resources of a cluster are insufficient, you can configure elastic scaling policies to increase nodes automatically. In the **Nodes** tab of the cluster details page, the nodes in the node group whose **Node Type** is **Task** are Task nodes. If only the **NodeManager (Yarn)** or **Supervisor (Storm)** role is deployed in a node group in addition to basic mandatory roles, this node group is a Task node group.

MRS cluster nodes support remote login. The following remote login methods are available:

- GUI login: Use the remote login function provided by the ECS console to log in to the Linux CLI of the cluster node.
- SSH login: Use a remote login tool like PuTTY to access Linux ECSs. You need to bind an EIP to the ECS first.

  Apply for and bind an EIP to a cluster node. For details, see **Virtual Private Cloud** > **User Guide** > **Elastic IP** > **Assigning an EIP and Binding It to an ECS** .

  You can log in to the ECS using a key or password. For details, see **Logging In to an MRS Cluster Node**.

4.1.3 MRS Cluster Node Specifications

Cluster Node Specifications

MRS supports host specifications determined by CPU, memory, and disk space.

Tenants share physical resources of ECSs, but can exclusively use resources of BMSs. BMSs can better meet your requirements for deploying key applications and services that require high performance (such as big data clusters and enterprise middleware systems) and a secure and reliable running environment.

If BMS specifications are used, Master node specifications cannot be scaled up.

NOTE

- More advanced instance specifications provide better data processing.
- Instance specifications may vary in different AZs. If no instance specifications in the current AZ can meet your requirements, switch to another AZ.
- If you select non-HDD disks for Core nodes, the disk types of Master and Core nodes are determined by the data disk.
- If **Sold out** appears next to an instance specification of a node, the node of this specification cannot be created. You can only create nodes of other specifications.
- For MRS 3.*x* or later, the memory allocated to Master and Core nodes must be greater than or equal to 64 GB.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Cluster Node Disk Types

Table 4-3 Disk types of MRS cluster nodes

Disk Role	Description
System disk	Storage type and space of the system disk on a node. Storage type can be any of the following: • **SAS**: high I/O • **SSD**: ultra-high I/O • **GPSSD**: general-purpose SSD • **ESSD**: extreme SSD
Data disk	Data disk storage space of a node. For more data storage, you can add disks when creating a cluster. A maximum of 10 disks can be added to each Core or Task node. • Data storage and computing are separated. Data is stored in OBS, which features low cost and unlimited storage capacity. The clusters can be deleted at any time in OBS. The computing performance is determined by OBS access performance and is lower than that of HDFS. This configuration is recommended if data computing is infrequent. • Data storage and computing are not separated. Data is stored in HDFS, which features high cost, high computing performance, and limited storage capacity. Before deleting clusters, you must export and store the data. This configuration is recommended if data computing is frequent. The storage type can be any of the following: • **SAS**: high I/O • **SSD**: ultra-high I/O • **GPSSD**: general-purpose SSD • **ESSD**: extreme SSD **NOTE** Adding nodes to an MRS cluster requires increasing the disk capacity of the management node (Master node). To ensure stable cluster running, set the disk capacity of the Master node to over 600 GB if the number of nodes is 300 and increase it to over 1 TB if the number of nodes reaches 500.

4.2 MRS Cluster Deployment

4.2.1 Overview

The analysis cluster, streaming cluster, and hybrid cluster provided by MRS use fixed templates to deploy cluster processes. Therefore, you cannot customize service processes on management nodes and control nodes.

If you want to customize the cluster deployment, set **Cluster Type** to **Custom** when creating a cluster. In this way, you can customize the deployment mode of

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

process instances on cluster nodes. **(Only MRS 3.*x* and later versions support the creation of custom topology clusters.)**

A custom cluster provides the following functions:

- Separated deployment of the management and control roles: The management role and control role are deployed on different Master nodes.
- Co-deployment of the management and control roles: The management and control roles are co-deployed on the Master node.
- Components are deployed separately to avoid resource contention.

Cluster Deployment Types

Table 4-4 MRS cluster deployment types

Common Node	Description	Node Range
Compact	The management role and control role are deployed on the Master node, and data instances are deployed in the same node group. This deployment mode applies to scenarios where the number of control nodes is less than 100, reducing costs.	• A minimum of 3 and a maximum of 11 Master nodes are required. • The total number of node groups is less than or equal to 10, and the total number of nodes in non-Master node groups is less than or equal to 10,000.
Co-deployment of the management role, control role, and data instances	This template is not recommended in the production environment or commercial environment. If management, control, and data nodes are deployed together, cluster performance and reliability may deteriorate. If there are enough nodes, deploy data nodes separately.	• A minimum of 3 and a maximum of 100 Master nodes are required. • The total node groups are no more than 10.
OMS-separate	The management role and control role are deployed on different Master nodes, and data instances are deployed in the same node group. This deployment mode is applicable to a cluster with 100 to 500 nodes and delivers better performance in high-concurrency load scenarios.	• A minimum of 5 and a maximum of 11 Master nodes are required. • The total number of node groups is less than or equal to 10, and the total number of nodes in non-Master node groups is less than or equal to 10,000.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common Node	Description	Node Range
Full-size	The management role and control role are deployed on different Master nodes, and data instances are deployed in different node groups. This deployment mode is applicable to a cluster with more than 500 nodes. Components can be deployed separately, which can be used for a larger cluster scale.	• A minimum of 9 and a maximum of 11 Master nodes are required. • The total number of node groups is less than or equal to 10, and the total number of nodes in non-Master node groups is less than or equal to 10,000.

Cluster Role Deployment Rules

The MRS system consists of multiple services in logical architecture. Each service contains one or more roles. Each role can deploy one or more instances.

- Service: A service is a kind of service capabilities provided by a component in a cluster. Each component in the cluster provides a service and is named by the service.
- Role: Roles are components of a service. Each service consists of one or more roles. Services are deployed on nodes (servers) through roles to ensure normal service running.
- Instance: An instance is formed when a service role is installed on a node. A service has one or more role instances.

For details about the deployment principles of each service, see **Overview**.

- Service A depending on service B: If service A is deployed in a cluster, service B must be deployed in advance. Service B provides basic capabilities for service A.
- Service A associated with service B: Service A exchanges data with service B during service running but does not depend on service B during deployment.
- Role A and role B deployed on the same server: If role A is deployed in a cluster, role B must also be deployed, and role A and role B must be deployed on the same node.

Table 4-5 MRS role deployment rules

Service	Dependency	Role	Role Deployment Suggestions
OMSServer	-	OMSServer	Deploy this role only on the Master node. Deployment on other nodes is not allowed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
ClickHouse	Depends on ZooKeeper.	CHS (ClickHouseServer)	Deploy this role on any node in the cluster. Valid role instance range: 2 to 256 (even numbers only).
		CLB (ClickHouseBalancer)	Deploy this role on any node in the cluster. Valid role instance range: 2 to 256.
Flink	• Depends on ZooKeeper. • Depends on KrbServer. • Depends on DBService. • Depends on Hadoop.	FR(FlinkResource)	Deploy this role on any node in the cluster. Valid role instance range: 1 to 10,000.
		FS(FlinkServer)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 2.
Flume	-	MS(MonitorServer)	Deploy this role only on the Master node. Valid role instance range: 1 to 2
		F(Flume)	Deploy this role on any node in the cluster. Valid role instance range: 1 to 10,000.
Hadoop	Depends on ZooKeeper.	NN(NameNode)	Deploy this role only on the Master node. Valid role instance range: 2.
		HFS (HttpFS)	Deploy this role only on the Master node. Valid role instance range: 0 to 10
		JN(JournalNode)	Deploy this role only on the Master node. Valid role instance range: 3 to 60, in increments of 2.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
		DN(DataNode)	Deploy this role on any node in the cluster. Valid role instance range: 3 to 10,000.
		RM(ResourceManager)	Deploy this role only on the Master node. Valid role instance range: 2.
		NM(NodeManager)	Deploy this role on any node in the cluster. Valid role instance range: 3 to 10,000.
		JHS(JobHistoryServer)	Deploy this role only on the Master node. Valid role instance range: 1 to 2
		TLS(TimelineServer)	Deploy this role only on the Master node. Valid role instance range: 0 to 1
HBase	Depends on Hadoop.	HM(HMaster)	Deploy this role only on the Master node. Valid role instance range: 2.
		TS(ThriftServer)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 10,000.
		RT(RESTServer)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 10,000.
		RS(RegionServer)	Deploy this role on any node in the cluster. Valid role instance range: 3 to 10,000.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
		TS1(Thrift1Server)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 10,000. If the Hue service is installed in a cluster and HBase needs to be used on the Hue web UI, install this instance for HBase.
HetuEngine	• Depends on Hadoop. • Depends on DBService. • Depends on Hive. • Depends on ZooKeeper. • Depends on KrbServer. • Depends on Yarn. • Depends on HDFS.	HSB(HSBroker)	Deploy this role on any node in the cluster. Valid role instance range: 2 to 50.
		HSC(HSConsole)	Deploy this role on any node in the cluster. Valid role instance range: 2.
		HSF(HSFabric)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 50.
		QAS (available for MRS 3.2.0-LTS.1 and later versions only)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 2.
Hive	• Depends on Hadoop. • Depends on DBService.	MS(MetaStore)	Deploy this role only on the Master node. Valid role instance range: 2 to 10.
		WH (WebHCat)	Deploy this role only on the Master node. Valid role instance range: 1 to 10.
		HS(HiveServer)	Deploy this role only on the Master node. Valid role instance range: 2 to 80.
Hue	Depends on DBService.	H(Hue)	Deploy this role only on the Master node. Valid role instance range: 2.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
Impala	• Depends on Hadoop. • Depends on Hive. • Depends on DBService. • Depends on ZooKeeper.	StateStore	Deploy this role only on the Master node. Valid role instance range: 1.
		Catalog	Deploy this role only on the Master node. Valid role instance range: 1.
		Impalad	Deploy this role on any node in the cluster. Valid role instance range: 1 to 10,000.
IoTDB	Depends on KrbServer.	ConfigNode (CN)	Deploy this role only on the Master node. Valid role instance range: 3 to 9, in increments of 2.
		IoTDBServer (IoTDBS)	Deploy this role on any node in the cluster. Valid role instance range: 3 to 256.
Kafka	Depends on ZooKeeper.	B(Broker)	Deploy this role on any node in the cluster. Valid role instance range: 3 to 10,000.
Kudu	-	KuduMaster	Deploy this role only on the Master node. Valid role instance range: 3 or 5.
		KuduTserver	Deploy this role on any node in the cluster. Valid role instance range: 3 to 10,000.
Loader	• Depends on Hadoop. • Depends on DBService.	LS(LoaderServer)	Deploy this role only on the Master node. Valid role instance range: 2.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
Oozie	• Depends on Hadoop. • Depends on DBService. • Depends on ZooKeeper.	O(oozie)	Deploy this role only on the Master node. Valid role instance range: 2.
Presto	Depends on Hive.	PCD(Coordinator)	Deploy this role only on the Master node. Valid role instance range: 2.
		PWK(Worker)	Deploy this role on any node in the cluster. Valid role instance range: 1 to 10,000.
Ranger	Depends on DBService.	RA(RangerAdmin)	Deploy this role only on the Master node. Valid role instance range: 1 to 2
		USC(UserSync)	Deploy this role only on the Master node. Valid role instance range: 1.
		TSC (TagSync)	Deploy this role on any node in the cluster. Valid role instance range: 0 to 1
Spark	• Depends on Hadoop. • Depends on Hive. • Depends on ZooKeeper.	JS(JDBCServer)	Deploy this role only on the Master node. Valid role instance range: 1 to 2
		JH(JobHistory)	Deploy this role only on the Master node. Valid role instance range: 1 to 2
		SR(SparkResource)	Deploy this role only on the Master node. Valid role instance range: 1 to 2

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
Spark2x	• Depends on Hadoop. • Depends on Hive. • Depends on ZooKeeper.	JS2X(JDBCServer2x)	Deploy this role only on the Master node. Valid role instance range: 2 to 10.
		JH2X(JobHistory2x)	Deploy this role only on the Master node. Valid role instance range: 2.
		SR2X(SparkResource2x)	Deploy this role only on the Master node. Valid role instance range: 2 to 50.
		IS2X(IndexServer2x)	(Optional) Deploy this role only on the Master node. Valid role instance range: 0 to 2, in increments of 2.
Sqoop	Depends on Hadoop.	SC(SqoopClient)	Deploy this role on any node in the cluster. Valid role instance range: 1 to 10,000.
Tez	• Depends on Hadoop. • Depends on DBService. • Depends on ZooKeeper.	TUI(TezUI)	Deploy this role only on the Master node. Valid role instance range: 1 to 2
ZooKeeper	-	QP(quorumpeer)	Deploy this role only on the Master node. Valid role instance range: 3 to 9, in increments of 2.
Doris (MRS 3.3.0-LTS and later)	-	FE	Deploy this role only on the Master node. Valid role instance range: 1 to 3, in increments of 2. **NOTICE** A single FE node applies only to the test or experiment environment. You are advised to deploy at least three FE nodes for the production environment.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
		BE	Deploy this role on any node in the cluster. Valid role instance range: 3 to 200.
		DBroker	(Optional) Deploy this role on any node in the cluster. Valid role instance range: 0 to 200.
JobGateway	• Depends on Hadoop. • Depends on DBService.	JS(JobServer)	Deploy this role only in the Master node group. Valid role instance range: 2 to 10.
		JB(JobBalancer)	Deploy this role in any node group. Valid role instance range: 2.
MemArtsCC	Depends on ZooKeeper.	CCWorker	Deploy this role on Core nodes. Valid role instance range: 2 to 9999.
		CCSidecar	Deploy this role on the Core nodes. Ensure that this role and the CCWorker are deployed on the same node. Valid role instance range: 2 to 9999.
MemArtsStore	Depends on ZooKeeper.	StoreWorker	Deploy this role on Core nodes. Valid role instance range: 3 to 5000.
		StoreSidecar	Do not deploy this role independently. Deploy it on the same nodes as the StoreWorker and StoreMaster roles.
		StoreClientProxy	(Optional) Deploy this role on any node in the cluster. Deploy this role on the same nodes as YARN NodeManager. Valid role instance range: 0 to 10,000.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Dependency	Role	Role Deployment Suggestions
		StoreMaster	Deploy this role in the Master node group. Do not deploy it on the same node as StoreWorker. Valid role instance range: 2 to 3.

4.2.2 Kerberos Authentication for MRS Clusters

The Hadoop community version has two authentication modes: Kerberos authentication for security and Simple authentication for regular use. During cluster setup, you can choose to enable or disable Kerberos authentication. Once the cluster is created, it cannot be modified.

Security Mode (Kerberos Authentication Enabled)

The MRS clusters in security mode use the Kerberos authentication protocol for security authentication. The Kerberos protocol supports mutual authentication between clients and servers. This eliminates the risks incurred by sending user credentials over the network for simulated authentication. In clusters, KrbServer provides the Kerberos authentication support.

Kerberos user object

In the Kerberos protocol, each user object is a principal. A complete principal consists of username and domain name. In O&M or application development scenarios, the user identity must be verified before a client connects to a server. Users for O&M and service operations are classified into human-machine and machine-machine users. The password of human-machine users is manually configured, while the password of machine-machine users is generated by the system randomly.

Kerberos authentication

Kerberos authentication supports two authentication modes: password authentication and keytab authentication. The default authentication validity period is 24 hours.

- Password authentication: User identity is verified by entering the correct password. This mode mainly used in O&M scenarios where human-machine users are used. The client command is **kinit** *Username*.
- Keytab authentication: Keytab files contain users' principal and encrypted credential information. When keytab files are used for authentication, the system automatically uses encrypted credential information to perform authentication and the user password does not need to be entered. This mode is mainly used in component application development scenarios where machine-machine users are used. Keytab authentication can also be configured using the **kinit** command.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common mode (Kerberos Authentication Disabled)

In normal clusters, MRS components use a native open source authentication mechanism, which is typically Simple authentication. If Simple authentication is used, authentication is automatically performed by a client user (for example, user **root**) by default when a client connects to a server. The authentication is imperceptible to the administrator or service user. In addition, when being executed, the client may even pretend to be any user (including **superuser**) by injecting **UserGroupInformation**. Cluster resource management and data control APIs are not authenticated on the server and are easily exploited and attacked by hackers.

It is best to use this for single-user scenarios in normal clusters, but make sure to tightly control network access permissions to keep the cluster secure.

- Deploy service applications on ECSs in the same VPC and subnet to avoid accessing the MRS cluster through the external network.
- Configure security group rules to strictly control the access scope. Do not configure access rules that allow **Any** or **0.0.0.0** for the inbound direction of MRS cluster ports.

4.2.3 Disk Partitions of an MRS Cluster Node

MRS clusters can be used immediately after being created. You do not need to plan disk partitions. **Table 4-6** describes the OS disk partitions of a created cluster node.

Table 4-6 OS disk partitions of an MRS cluster node

Partition Type	Partition Directory	Capacity	Usage
OS partition	/	220 GB	OS root partition directory and program storage directory, including all directories (except the directories specified below)
	/tmp	10 GB	Directory for storing temporary files
	/var	10 GB	OS runtime directory
	/var/log	The remaining space of the OS disk is allocated to the **/var/log** partition.	Default directory for storing MRS cluster logs

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Partition Type	Partition Directory	Capacity	Usage
	/srv/BigData	60 GB	Data directory of MRS Manager, which stores data such as ldapData, Manager, and metric_agent, and provides mount points for component data directories.

NOTE

This section applies only to MRS 3.*x* or later. The OS disk of MRS 3.3.1-LTS or later is not partitioned and has only the **/** directory.

After an MRS node is created, the non-OS disks of the node are mounted to the **/srv/BigData/data*N*** directory. For example, if the node has four data disks, the disk mount directories are **/srv/BigData/data1**, **/srv/BigData/data2**, **/srv/BigData/data3**, and **/srv/BigData/data4**.

The metadata and data directories of each component deployed on the node are allocated to different disk partition directories based on certain mapping rules. **Table 4-7** describes the data directories of each component.

Table 4-7 Non-OS disk partitions of an MRS cluster node

Partition Type	Disk Partition Mount Directory	Data Directory	Usage
Metadata partition	/srv/BigData/data1	dbdata_om	Directory for storing OMS database data. If FusionInsight Manager is to be installed on two nodes, both OMS nodes contain this directory.
		LocalBackup	If **LocalDir** is selected for backing up cluster data, the backup data is stored in the directory by default. If FusionInsight Manager is to be installed on two nodes, both nodes contain this directory.
		doris/fe	Directory for storing Doris metadata.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Partition Type	Disk Partition Mount Directory	Data Directory	Usage
		• kudu/master/metadata • kudu/master/wal	Directory for storing Kudu Master metadata and WAL data.
	/srv/BigData/data2	journalnode	Node where the JournalNode role of HDFS is located, which stores the JournalNode metadata of HDFS.
		dbdata_service	Node where the DBServer role of DBService is located, and the database directory of DBService
		iotdb/iotdbserver	Directory for storing IoTDB metadata
		iotdb/confignode	Directory for storing metadata of the IoTDB ConfigNode role
	/srv/BigData/data3	namenode	Node where the NameNode role of HDFS is located, which stores NameNode data.
		iotdb/iotdbserver	Directory for storing IoTDBServer log data
	/srv/BigData/data4	zookeeper	Node where the QuorumPeer role of ZooKeeper is located, which stores ZooKeeper data.
		hetuengine/qas	Node where the QAS role of HetuEngine is located, which stores QAS data.
Service data partition	/srv/BigData/data*N*	• dn • nm	Directory for storing DataNode data and intermediate MapReduce data
		kafka-logs	Directory for storing Kafka broker data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Partition Type	Disk Partition Mount Directory	Data Directory	Usage
		clickhouse clickhouse_path	Directory for storing ClickHouse database data. The **clickhouse_path** directory storing ClickHouse metadata exists only in the **data1** directory.
		iotdb/ iotdbserver	Directory for storing IoTDB service data.
		memarts	Directory for storing MemArtsCC cache data.
		storeworker	Directory for storing MemArtsStore shuffle data.
		doris/be	Directory for storing Doris database data.
		kudu	Directory for storing Kudu data.
		impala	Directory for storing Impala data.

- The metadata partition directory uses a maximum of four disks (data1 to data4). The metadata directories are mapped to directories from **/srv/BigData/data1** to **/srv/BigData/data4** in sequence according to **Table 4-7**. If only three data disks are mounted to the current node, the directories under data4 and data2 are combined. If only two data disks are mounted, the directories under data3 and data1 are combined, and those under data4 and data2 are combined.

  For example, if the ZooKeeper node has four data disks, the ZooKeeper data directory is **/srv/BigData/data4/zookeeper**. If the node has only three data disks, the ZooKeeper data directory is **/srv/BigData/data2/zookeeper**.
- The mapping rules of service data directories are as follows:

  For HDFS, Kafka, ClickHouse, and IoTDB, mount points that comply with the **/srv/BigData/data*N*** directory are automatically identified as data directories based on the number of mounted disks on the current node.

  For example, if disks are mounted to the **/srv/BigData/data1** to **/srv/BigData/data3** directories, the DataNode data directories are **/srv/BigData/data1/dn**, **/srv/BigData/data2/dn**, and **/srv/BigData/data3/dn**, and the Kafka data directories are **/srv/BigData/data1/kafka-logs**, **/srv/BigData/data2/kafka-logs**, and **/srv/BigData/data3/kafka-logs**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

5 Buying MRS Clusters

5.1 Quickly Creating an MRS Cluster

MRS consists of multiple big data components, and you can select the cluster type that best fits your service requirements, data types, reliability expectations, and resource budget.

This section uses an HBase query cluster as an example to describe how to quickly create an MRS cluster. An HBase cluster uses Hadoop and HBase components to provide a column-oriented distributed cloud storage system featuring enhanced reliability, great performance, and elastic scalability. It applies to the storage and distributed computing of massive amounts of data. You can use HBase to build a storage system capable of storing TB- or even PB-level data. With HBase, you can filter and analyze data with ease and get responses in milliseconds, rapidly mining data value.

Procedure

Step 1 Log in to the MRS console.

Step 2 Click **Create Cluster**. The page for creating a cluster is displayed.

Step 3 On the displayed page, click the **Quick Config** tab.

Step 4 Configure basic cluster information by referring to the following table.

Table 5-1 MRS cluster parameters

Parameter	Description	Example Value
Region	Region where the resource to be created is located. Resources located in different regions cannot communicate with each other using an internal network. To improve access speed and reduce network latency, choose a region that is close to your own.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Cluster Name	MRS cluster name. You can use the default name. However, you are advised to include a project name abbreviation or date for consolidated memory and easy distinguishing. After a cluster is created, you can change the cluster name in the cluster list.	mrs-test
Cluster Type	Select a proper MRS cluster type based on service requirements. • **Analysis cluster**: It is used for offline data analysis and comprises data analysis tools like Hadoop, Spark, HBase, Hive, Flink, Oozie, and Tez. • **Streaming cluster**: It processes streaming data to quickly analyze real-time data sources, and it mainly includes streaming data processing tools such as Kafka and Flume. • **Hybrid cluster**: It is suitable for both offline data analysis and stream processing. • **Custom**: You can select from a variety of components that are supported by the corresponding version of the MRS cluster. • **Doris cluster**: Doris is a high-performance and real-time analytical database for report analysis, ad hoc queries, and data lake federated query acceleration.	Custom

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Version Type	MRS provides two cluster version types: **LTS** and **Normal**. Different version types provide different components. You can select a version type as required. • **LTS**: employs MRS's own components to provide highly reliable clusters with strong DR capabilities, making long-term support and evolution possible. • **Normal**: integrates MRS's mature and stable features and functions with open-source capabilities, offering high performance and stability.	LTS
Cluster Version	Version of the MRS cluster. Different versions may contain different open-source component versions and functions. You are advised to select the latest version.	MRS 3.2.0-LTS.1
Component	Cluster templates containing preset components you will need for your business. **NOTE** Hadoop components include HDFS, YARN, and MapReduce. Components used within a cluster, such as DBService, KrbServer, and LdapServer, are not displayed in the component list during cluster creation.	HBase Query Cluster
AZ	AZ to which resources to be created belong in the current region. An AZ is a physical area where resources use independent power supply and networks.	AZ1
VPC	VPC to which the MRS cluster nodes belong. If no VPC is available, click **View VPC** to open the network console and create a VPC.	-
Subnet	Subnet in the VPC. If no subnet is available, click **View Subnet** to open the network console and create a subnet.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Cluster Node	Specifications and quantity of nodes in an MRS cluster.	Set the number of cluster nodes and node specifications as required.
Kerberos Authentication	Whether to enable Kerberos authentication for each component in the MRS cluster. If Kerberos authentication is enabled, users can access component resources only after being authenticated. This option cannot be changed after you create a cluster.	Kerberos authentication enabled
Username	Default user for logging in to MRS Manager and nodes in the MRS cluster. User **admin** is used to log in to MRS Manager, while user **root** is used to log in to the OS of nodes in the cluster.	-
Password/ Confirm Password	Passwords required for the **root** and **admin** users. The passwords are user-defined and must be kept secure.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Kerberos Encryption Type	Encryption algorithm and method used by Kerberos. (This parameter is available in MRS 3.3.1-LTS or later. In earlier versions, the default Kerberos encryption type is **aes256-sha1,aes128-sha1**.) ● **aes256-sha1,aes128-sha1**: indicates that the encryption algorithm and mode are AES256-CTS-HMAC-SHA1-96 AES128-CTS-HMAC-SHA1-96. ● **aes256-sha2,aes128-sha2**: indicates that the encryption algorithm and mode are AES256-CTS-HMAC-SHA384-192 AES128-CTS-HMAC-SHA256-128. **When setting this parameter, note the following:** ● If the encryption types of two clusters are different, mutual trust cannot be set up across the FusionInsight Manager clusters. ● If you are using an external Java program to connect to the MRS cluster for Kerberos authentication and the Kerberos encryption type is **aes256-sha2,aes128-sha2**, the external Java program must use JDK 11 (or BiSheng JDK 1.8.0_392) or later. ● To use an instance created in DataArts Studio to connect to the MRS cluster, set the Kerberos encryption type to **aes256-sha1,aes128-sha1**.	aes256-sha2,aes128-sha2
Enterprise Project	An enterprise project facilitates project-level management and grouping of cloud resources and users. You can select the system-defined enterprise project **default** or create a new one.	default

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Required Duration	Cluster subscription duration. This parameter is valid in Yearly/Monthly billing mode. The minimum cluster duration is 1 month and the maximum available cluster duration is 1 year. If **Auto-renew** is selected, monthly subscriptions are automatically renewed every month and yearly subscriptions are automatically renewed every year.	-
Secure Communications	To allow the MRS console to access big data components in your VPC, you need to enable relevant security group rules to allow traffic to pass. For details, see **Configuring Secure Communication Authorization for an MRS Cluster**.	Select the check box.

Step 5 Click **Create Now**.

If Kerberos authentication is enabled for a cluster, check whether Kerberos authentication is required. If yes, click **Continue**. If no, click **Back** to disable Kerberos authentication and then create a cluster. This option cannot be changed after you create a cluster.

Step 6 Click **Back to Cluster List** to view the cluster status. Click **Access Cluster** to view cluster details.

For details about cluster status during creation, see the description of the status parameters in **Table 8-8**.

It takes some time to create a cluster. The initial status of the cluster is **Starting**. After the cluster has been created successfully, the cluster status becomes **Running**.

On the MRS console, a maximum of 10 clusters can be created concurrently, and a maximum of 100 clusters can be managed.

----End

FAQs

The following lists the questions frequently asked during MRS cluster creation and their possible causes:

How Do I View Details About Failed MRS Cluster Creation Tasks

If a cluster fails to be created, the failed task will be managed on the **Manage Failed Tasks** page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

On the cluster list page, click  to switch to the **Manage Failed Tasks** page. In the **Task Status** column, hover the cursor over the task status to view the failure cause.

Table 5-2 lists the error codes of MRS cluster creation failures.

Table 5-2 Error codes

Error Code	Description
MRS.101	Insufficient quota to meet your request. Contact customer service to increase the quota.
MRS.102	The token cannot be null or invalid. Try again later or contact customer service.
MRS.103	Invalid request. Try again later or contact customer service.
MRS.104	Insufficient resources. Try again later or contact customer service.
MRS.105	Insufficient IP addresses in the existing subnet. Try again later or contact customer service.
MRS.201	Failed due to an ECS error. Try again later or contact customer service.
MRS.202	Failed due to an IAM error. Try again later or contact customer service.
MRS.203	Failed due to a VPC error. Try again later or contact customer service.
MRS.400	MRS system error. Try again later or contact customer service.

Helpful Links

- After an MRS cluster is created, you can install the MRS cluster client to connect and use the components in the cluster. For details, see **Installing an MRS Cluster Client**.
- For details about how to submit and run jobs in an MRS cluster, see **Submitting an MRS Job**.
- You can log in to FusionInsight Manager of a cluster to view cluster resource status, alarms, and other information. For details, see **Accessing MRS Manager**.

5.2 Creating a Custom MRS Cluster

This section describes how to create an MRS cluster on the MRS console.

MRS consists of multiple big data components, and you can select the cluster type that best fits your service requirements, data types, reliability expectations, and resource budget.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can **quickly buy** a cluster using the preset cluster template or select the component list and advanced settings to manually buy a cluster.

On the MRS console, a maximum of 10 clusters can be created concurrently, and a maximum of 100 clusters can be managed.

Prerequisites

- You have performed MRS authorization.
- You have learned the MRS cluster planning and determined the resource scale of clusters. For details, see **MRS Cluster Planning**.

Creating a Custom MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Click **Create Cluster**. The page for creating a cluster is displayed.

Step 3 Click the **Custom Config** tab.

Pay attention to quota reminders when creating a cluster. If the resource quota is insufficient, request a quota increase as prompted before creating the cluster.

Step 4 In the **Basic Configuration** area, set basic information about the MRS cluster.

- **Region**: Region where the resource to be created is located. Resources in different regions cannot communicate with each other over the intranet. To improve access speed and reduce network latency, select the region nearest to you.
- **Billing Mode**: Billing mode of a cluster. MRS provides two billing modes: yearly/monthly and pay-per-use.

Step 5 Configure the MRS cluster information as prompted.

When creating a cluster, you need to configure the cluster version, network, node specifications and quantity, and other advanced settings.

- **MRS Cluster Configuration**
- **MRS Cluster Network Configuration**
- **MRS Cluster Node Configuration**
- **Other MRS Cluster Configuration Parameters**

Step 6 After the cluster information is configured, click **Create Now**.

Step 7 Click **Back to MRS Cluster List** to view the cluster status.

Wait for the cluster creation to complete. The initial status of the cluster is **Starting**. After the cluster is created, the cluster status becomes **Running**.

For details about cluster status during creation, see the description of the status parameters in **Table 8-8** of **Checking the Running Status of an MRS Cluster**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS Cluster Configuration

Table 5-3 MRS cluster parameters

Parameter	Description	Example Value
Cluster Name	The cluster name must be unique. It can contain 1 to 64 characters. Only letters, digits, hyphens (-), and underscores (_) are allowed.	mrs_test
Cluster Type	Select a proper MRS cluster type based on service requirements. • **Analysis cluster**: It is used for offline data analysis and comprises data analysis tools like Hadoop, Spark, HBase, Hive, Flink, Oozie, and Tez. • **Streaming cluster**: It processes streaming data to quickly analyze real-time data sources, and it mainly includes streaming data processing tools such as Kafka and Flume. • **Hybrid cluster**: It is suitable for both offline data analysis and stream processing. • **Custom**: You can select from a variety of components that are supported by the corresponding version of the MRS cluster. • **Doris cluster**: Doris is a high-performance and real-time analytical database for report analysis, ad hoc queries, and data lake federated query acceleration.	Custom
Version Type	MRS provides two cluster version types: **LTS** and **Normal**. Different version types provide different components. You can select a version type as required. • **LTS**: employs MRS's own components to provide highly reliable clusters with strong DR capabilities, making long-term support and evolution possible. • **Normal**: integrates MRS's mature and stable features and functions with open-source capabilities, offering high performance and stability.	LTS
Cluster Version	Version of the MRS cluster. Different versions may contain different open-source component versions and functions. You are advised to select the latest version.	MRS 3.2.0-LTS.1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Component	Select the components to be deployed in the MRS cluster. You can change components based on your needs. For some clusters, components cannot be added after creation. **NOTE** Hadoop components include HDFS, YARN, and MapReduce. Components used within a cluster, such as DBService, KrbServer, and LdapServer, are not displayed in the component list during cluster creation.	-
Metadata	Whether to use external data sources to store Hive and Ranger metadata of the cluster. • **Local**: Metadata is stored in the local cluster. • **External data connection**: Metadata of external data sources is used. If the cluster is abnormal or deleted, metadata is not affected. This mode applies to scenarios where storage and compute are decoupled. After a cluster is created, you can manually store the component metadata to external data sources. For details, see **Managing MRS Cluster Metadata**.	Local
Component Port	Default communication port policy of each component in the MRS cluster. The LTS cluster supports the configuration. • **Open source**: Use the port provided by the open-source component. • **Custom**: Customize a port for the component. For details about the differences between default open-source port and default custom port, see **Common Ports for MRS Cluster Services**.	Open source

MRS Cluster Network Configuration

Table 5-4 Network configuration parameters

Parameter	Description	Example Value
AZ	AZ to which resources to be created belong in the current region. An AZ is a physical area where resources use independent power supply and networks.	AZ1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
VPC	VPC to which the MRS cluster nodes belong. If no VPC is available, click **View VPC** to open the network console and create a VPC.	-
Subnet	Subnet in the VPC. If no subnet is available, click **View Subnet** to open the network console and create a subnet. • A subnet provides dedicated network resources that are logically isolated from other networks for network security. For details about how to configure network ACL outbound rules, see **How Do I Configure a Network ACL Outbound Rule?**. • The number of IP addresses required by creating an MRS cluster depends on the number of cluster nodes and selected components, but not the cluster type. • In MRS, IP addresses are automatically assigned to clusters during cluster creation basically based on the following formula: Quantity of IP addresses = Number of cluster nodes + 2 (one for Manager; one for the DB). In addition, if the Hadoop, Hue, Sqoop, and Presto or Loader and Presto components are selected during cluster deployment, one IP address is added for each component. To create a ClickHouse cluster separately, the number of IP addresses required is calculated as follows: Number of IP addresses = Number of cluster nodes + 1 (for FusionInsight Manager).	-
Security Group	A security group is a set of ECS access rules. It provides access policies for ECSs that have the same security protection requirements and are mutually trusted in a VPC. When you create an MRS cluster, a security group is automatically created by default. You can also select an existing security group from the drop-down list. • When you select a security group created by yourself, ensure that there is an inbound rule that allows traffic from all protocols, all ports, and the trusted accessible IP address range. Do not use **0.0.0.0/0** as a source address. Otherwise, security risks may occur. If you do not know the trusted accessible IP address range, select **Auto create**.	Auto create

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
EIP	After binding an EIP to an MRS cluster, you can use the EIP to access the Manager web UI of the cluster. • When creating a cluster, you can select an available EIP from the drop-down list and bind it. If no EIP is available in the drop-down list, click **Create EIPs** to create an EIP. • The EIP must be in the same region as the cluster.	Bind later

MRS Cluster Node Configuration

Table 5-5 Cluster node information

Parameter	Description	Example Value
Common Node Configurations	This parameter is available only when **Cluster Type** is set to **Custom**. Value options include **Compact**, **Full-size**, and **OMS-separate**. For details, see **Cluster Deployment Types**.	Compact
Node Group	Name of the node group in the cluster. • The name of the Master node group is fixed to **master_node_default_group**. • The system automatically creates a Core node group based on the components contained in the cluster. For example, if you select the ClickHouse component, the system adds the ClickHouse node group and deploys the **ClickHouseServer** role in the node group by default. • If **Cluster Type** is set to **Custom**, you can customize the names of other node groups. • If the data volume does not change much but the service processing capability changes greatly, manually add a **Task** node group.	node_group_1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Node Type	If **Cluster Type** is set to **Custom**, you can select the node type of a non-Master node group. If the node group type is set to **Task**, only the **NodeManager** role, and certain mandatory node roles, can be deployed in the node group. Nodes in a cluster can be classified into Master, Core, and Task nodes. • Master node: MRS cluster management node. OMSServer is deployed on the node to manage and monitor the cluster. • Core node: MRS cluster work node, which processes and analyzes data and stores process data. • Task node: Compute node. When the compute resources in a cluster are insufficient, you can configure elastic scaling policies to increase nodes automatically. For details about node types, see **MRS Cluster Node Types**.	Core
Instance Specifications	Select the instance specifications of the MRS cluster node. You can click to adjust the specifications. • For details about the MRS cluster node specifications, see **MRS Cluster Node Specifications**. • More advanced instance specifications provide better data processing. • Instance specifications may vary in different AZs. If no instance specifications in the current AZ can meet your requirements, switch to another AZ. • If you select non-HDD disks for Core nodes, the disk types of Master and Core nodes are determined by **Data Disk**. • If **Sold out** appears next to an instance specification of a node, the node of this specification cannot be created. You can only create nodes of other specifications. • For MRS 3.*x* or later, the memory allocated to Master and Core nodes must be greater than or equal to 64 GB.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
System Disk	You can adjust the storage type and space of the system disk on a node as required. For details about the MRS cluster storage, see **Cluster Node Disk Types**.	-
Data Disk	Storage type and space of data disks on a node. To increase the data storage capacity, you can add disks during cluster creation. A maximum of 10 disks can be added to each **Core** or **Task** node. For more information about MRS cluster storage, see **Cluster Node Disk Types**.	-
Instance Count	Number of Master and Core nodes. • Master Node Groups: The number of Master instances ranges from 3 to 9. • There must be at least one Core node and the total number of Core and Task nodes must not exceed 10,000. Click ⊕ to add a node group, click ✎ to modify the node instance specifications, and click 🗑 to delete the added node group. A small number of nodes may cause clusters to run slowly while a large number of nodes may be unnecessarily costly. Set an appropriate value based on data to be processed.	-
LVM	This parameter is valid when a streaming Core node is created only. Click this parameter to enable or disable the disk LVM management function. This parameter is not available in MRS 3.*x* and later versions. If LVM is enabled, all disks on a node are mounted as logical volumes. This delivers more proper disk planning to avoid data skew, thereby improving system stability.	Disable this function.
Topology Adjustment	If **Cluster Type** is set to **Custom**, you can adjust the deployment of each component in the cluster in the node group. Set **Topology Adjustment** to **Enable** and adjust the instance deployment mode based on service requirements. For details, see **Cluster Role Deployment Rules**.	Disable this function.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Other MRS Cluster Configuration Parameters

Table 5-6 Other configuration parameters

Parameter	Description	Example Value
Kerberos Authentication	Whether to enable Kerberos authentication for each component in the MRS cluster. If Kerberos authentication is enabled, users can access component resources only after being authenticated. This option cannot be changed after you create a cluster.	Kerberos authentication enabled
Username	Name of the administrator of Manager. **admin** is used by default.	admin
Password/ Confirm Password	Password of the Manager administrator **admin**. Keep the password secure. ● Must contain 8 to 26 characters. ● Must contain at least four of the following: – Lowercase letters – Uppercase letters – Digits – At least one of the following special characters: `~!@#$%^&*()-_=+|[{}];:',<.>/? ● Cannot be the same as the username or the username spelled backwards.	-
Login Mode	Method for logging in to a node in the MRS cluster. ● Password You can log in to the node as user **root** using a password. You need to customize the password of user **root**. ● Key Pair Select a key pair from the drop-down list. Select "I acknowledge that I have obtained private key file *SSHkey-xxx* and that without this file I will not be able to log in to my ECS." Click **View Key Pair** to create or import a key pair, and then obtain the private key file.	Password

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Kerberos Encryption Type	Encryption algorithm and method used by Kerberos. (This parameter is supported in MRS 3.3.1-LTS and later versions. In earlier versions, the default Kerberos encryption type is **aes256-sha1,aes128-sha1**.) • **aes256-sha1,aes128-sha1**: indicates that the encryption algorithm and mode are AES256-CTS-HMAC-SHA1-96 AES128-CTS-HMAC-SHA1-96. • **aes256-sha2,aes128-sha2**: indicates that the encryption algorithm and mode are AES256-CTS-HMAC-SHA384-192 AES128-CTS-HMAC-SHA256-128. **When setting this parameter, note the following:** • **After a cluster is created, the Kerberos encryption type cannot be modified. Configure the encryption type based on your service requirements.** • If the encryption types of two clusters are different, mutual trust cannot be set up across the FusionInsight Manager clusters. • If you are using an external Java program to connect to the MRS cluster for Kerberos authentication and the Kerberos encryption type is **aes256-sha2,aes128-sha2**, the external Java program must use JDK 11 (or BiSheng JDK 1.8.0_392) or later. • To use an instance created in DataArts Studio to connect to the MRS cluster, set the Kerberos encryption type to **aes256-sha1,aes128-sha1**.	aes256-sha2,aes128-sha2
Set Advanced Options	Advanced function parameters of the MRS cluster. For details, see **Table 5-7**.	-
Enterprise Project	Select the enterprise project to which the cluster belongs. To use an enterprise project, create one on the **Enterprise** > **Project Management** page. The **Enterprise Management** console is designed for resource management. It helps you manage cloud-based personnel, resources, permissions, and finance in a hierarchical manner, such as management of companies, departments, and projects.	default

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Secure Communications	To allow the MRS console to access big data components in your VPC, you need to enable relevant security group rules to allow traffic to pass. For details, see **Configuring Secure Communication Authorization for an MRS Cluster**.	Select the check box.

Table 5-7 MRS cluster advanced configuration

Parameter	Description	Example Value
Hostname Prefix	Enter the prefix for the computer hostname of an ECS in the cluster.	-
Cryptographic Algorithm	Algorithm used for encrypting and decrypting passwords in the cluster system. • **International**: general cryptographic algorithm • **Chinese**: SM series cryptographic algorithms are compatible with general cryptographic algorithms.	International
Tag	It is recommended that you use the tag function of TMS to add the same tag to different cloud resources. For details, see **Adding a Tag to an MRS Cluster/Node**.	-
Auto Scaling	Configure an auto scaling policy for the Task node group after the cluster is created. For details, see **MRS Task Node Auto Scaling**.	-
Bootstrap Action	Bootstrap actions run scripts on specified nodes in a cluster to install third-party software and modify the cluster's running environment. For details, see **Adding MRS Node Bootstrap Actions and Installing Third-Party Software**.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Agency	When an agency is bound to an ECS or BMS, it can manage designated resources in the cluster. It is essential to consider service requirements before deciding to configure an agency. For example, you can configure an ECS agency to automatically obtain the AK/SK to access OBS. For details, see **Interconnecting an MRS Cluster with OBS Using an IAM Agency**. To bind an agency to an MRS cluster, you need to create an IAM agency with required permissions in advance. By default, the system generates an **MRS_ECS_DEFAULT_AGENCY** agency. This agency has the **OBSOperateAccess** permission and the **CESFullAccess** (for users who have enabled fine-grained policies), CES Administrator, and KMS Administrator permissions in the region where the cluster is located.	Bind later
Metric Sharing	Monitoring metrics of big data components are collected. If a fault occurs when you use a cluster, share the monitoring metrics with technical support for troubleshooting.	Disable

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
System Disk Encryption	Whether to encrypt data in the system disk mounted to the MRS cluster node. This function is disabled by default. This parameter is available only in MRS 3.3.0-LTS or later. The Key Management Service (KMS) in Data Encryption Workshop (DEW) provides keys for system disk encryption. No additional key management infrastructure needs to be built or maintained. To enable this function, you must have the **Security Administrator** and **KMS Administrator** permissions. Configure the following parameters to enable this function: • **Data Disk Key ID**: Key ID of the selected key name. • **Data Disk Key Name**: Select the name of the key used to encrypt the data disk. By default, the default master key named **evs/default** is selected. You can select another master key from the drop-down list. If cloud disks are encrypted using a master key and the key is then disabled or scheduled for deletion, the disks can no longer be read from or written to, and data on these disks may never be recovered. Exercise caution when performing relevant operations. Click **View the Key List** to enter a page where you can create and manage keys.	Disable

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Data Disk Encryption	Whether to encrypt data in the data disk mounted to the MRS cluster node. This function is disabled by default. The Key Management Service (KMS) in Data Encryption Workshop (DEW) provides keys for system disk encryption. No additional key management infrastructure needs to be built or maintained. To enable this function, you must have the **Security Administrator** and **KMS Administrator** permissions. Configure the following parameters to enable this function: • **Data Disk Key ID**: Key ID of the selected key name. • **Data Disk Key Name**: Select the name of the key used to encrypt the data disk. By default, the default master key named **evs/default** is selected. You can select another master key from the drop-down list. If cloud disks are encrypted using a master key and the key is then disabled or scheduled for deletion, the disks can no longer be read from or written to, and data on these disks may never be restored. Exercise caution when performing this operation. Click **View the Key List** to enter a page where you can create and manage keys.	Disable
Alarm	If the alarm function is enabled, the cluster maintenance personnel can be notified in a timely manner to locate faults when the cluster runs abnormally or the system is faulty. To send alarm messages, you need to enable the Simple Message Notification (SMN) service, set a notification rule, and bind the rule to an SMN topic. • **Rule Name**: Name of the rule for sending alarm messages. The value can contain only digits, letters, hyphens (-), and underscores (_). • **Topic Name**: Select an existing SMN topic or click **Create Topic** to create a topic. To receive messages published to a topic, you need to add a subscriber to the topic. For details, see **Configuring Notifications for MRS Cluster Alarms and Events**.	Enable
Install UniAgent	Install UniAgent on MRS cluster nodes to simplify plug-in management, provide AOM instructions, and enable script delivery and execution.	No

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Logging	Whether to collect logs when cluster creation fails. After the logging function is enabled, system logs and component run logs are automatically collected and saved to the OBS file system in scenarios such as cluster creation failures and scale-out or scale-in failures for O&M personnel to quickly locate faults. The log information is retained for a maximum of seven days.	Disable

FAQs

The following lists the questions frequently asked during MRS cluster creation and their possible causes:

How Do I View Details About Failed MRS Cluster Creation Tasks

If a cluster fails to be created, the failed task will be managed on the **Manage Failed Tasks** page.

On the cluster list page, click  to switch to the **Manage Failed Tasks** page. In the **Task Status** column, hover the cursor over the task status to view the failure cause.

Table 5-8 lists the error codes of MRS cluster creation failures.

Table 5-8 Error codes

Error Code	Description
MRS.101	Insufficient quota to meet your request. Contact customer service to increase the quota.
MRS.102	The token cannot be null or invalid. Try again later or contact customer service.
MRS.103	Invalid request. Try again later or contact customer service.
MRS.104	Insufficient resources. Try again later or contact customer service.
MRS.105	Insufficient IP addresses in the existing subnet. Try again later or contact customer service.
MRS.201	Failed due to an ECS error. Try again later or contact customer service.
MRS.202	Failed due to an IAM error. Try again later or contact customer service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Error Code	Description
MRS.203	Failed due to a VPC error. Try again later or contact customer service.
MRS.400	MRS system error. Try again later or contact customer service.

Helpful Links

- After an MRS cluster is created, you can install the MRS cluster client to connect and use the components in the cluster. For details, see **Installing an MRS Cluster Client**.
- For details about how to submit and run jobs in an MRS cluster, see **Submitting an MRS Job**.
- You can log in to FusionInsight Manager of a cluster to view cluster resource status, alarms, and other information. For details, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

6 Installing an MRS Cluster Client

6.1 Installing a Client (MRS 3.*x*)

An MRS client is a tool used to interact with clusters. You can use the client to submit jobs, manage cluster resources, and view job status. The client works closely with the cluster together to meet your requirements for large-scale data processing and analytics. The client sends user job requests to the cluster. The cluster then executes jobs efficiently and returns results and status to the client.

You need to install a cluster client to connect to the component server in the cluster and perform tasks such as component connection and job submission. You can install the cluster client on a node within the cluster or on a node outside of it.

To ensure that certain features function properly, it is important to reinstall the client after modifying the server configuration of a component in the cluster. This ensures that the client version matches the server version.

Notes and Constraints

- This section describes how to install the MRS 3.*x* cluster client. For MRS 2.*x* and earlier versions, see **Installing a Client (MRS 2.x or Earlier)**.
- If the cluster contains the Flume component, install the Flume client separately.

Prerequisites

- If the node where the client is to be installed is outside the MRS cluster, the node must be able to communicate with the nodes in the MRS cluster. Otherwise, client installation will fail.
- The node where the client is to be installed must have the NTP service enabled and synchronized time with the MRS cluster server. Otherwise, client installation will fail.
- You can install the client as user **root** or any OS user. The user must have the operation permission on the client file storage directory and installation directory. For example, the permission on the two directories is **755**. Users **root** and **omm** have the required permission by default. You can enter the required directory and run the **ll** command to view the permission.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- When you install the client as a user other than **omm** or **root**, and the **/var/tmp/patch** directory already exists, you have changed the permission for the directory to **777** and changed the permission for the logs in the directory to **666**. The command for changing the permission is **chmod** *[who] operator [permission] filename*.

Installing a Client on a Node Inside a Cluster

Step 1 Obtain the client software package.

1. Log in to the MRS console.
2. In the navigation pane, choose **Active Clusters**. Click the target cluster name to access the cluster details page.
3. Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP**, configure an EIP, select the check box, and click **OK**.
   - If no EIP is available, click **Manage EIPs** to create one.
   - For details about how to access FusionInsight Manager (also called MRS Manager), see **Accessing MRS Manager**.
4. On the displayed page, enter the username and password, and click **Log In**.
5. Download the client on FusionInsight Manager.
   - Versions earlier than MRS 3.3.0: Click **Cluster**. On the **Dashboard** page, choose **More** > **Download Client**, set the parameters in **Table 6-1**, and click **OK**.
   - MRS 3.3.0 or later: On the FusionInsight Manager homepage, click **Download Client**, set the parameters in **Table 6-1**, and click **OK**.

   The client software package downloaded from the FusionInsight Manager homepage contains the clients of all services (excluding Flume) in the cluster. To download the client of a single service, choose **Cluster** > **Services** > *Service name*, click **More**, and select **Download Client**.

Table 6-1 Client download parameters

Parameter	Description	Example Value
Select Client Type	– **Complete Client**: contains the complete client software package and configuration files. Generally, this option is selected. – **Configuration Files Only**: downloads only client configuration files in the scenario where the administrator modifies the component server configuration on FusionInsight Manager after the complete client is downloaded and installed in an application development task, and developers need to update client configuration files.	Complete Client

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Select Platform Type	**The client type must match the architecture of the node where the client is to be installed. Otherwise, the installation fails.** **For clusters of the LTS version, only the client software package whose type is the same as that of FusionInsight Manager can be downloaded.** – **x86_64**: indicates the client software package that can be deployed on an x86 platform. – **aarch64**: indicates the client software package that can be deployed on a Kunpeng server.	x86_64
Save to Path	The path for storing the client software package on the active OMS node – Select **Save to Path**: Customize the path for storing the client software package on the active OMS node. User **omm** must have the read, write, and execute permissions on the path. If the path is not changed, the client file generated is saved in the **/tmp/FusionInsight-Client** directory on the active OMS node in the cluster by default. – Not to select **Save to Path**: The generated client file is automatically downloaded and saved to the local host. Before installing the client, you need to upload the file to a specified directory on the target node.	Select **Save to Path**
Select Download Location	The location for saving the downloaded client file **(only supported by MRS 3.3.1 or later)**. – **Server**: Download the file to the active OMS node of the cluster. – **Browser**: Download the file to the local computer. – **Remote Node:** Download the file to a node other than the active OMS node. If you select a remote location, configure the required parameters by referring to **Table 6-2**.	Server

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Save to Path	The path for storing the client software package **(only supported by MRS 3.3.1 or later)**. – If the download location is set to **Browser**: You do not need to set this parameter. – If the download location is set to **Server**: ▪ The default directory is **/tmp/FusionInsight-Client** on the active OMS node. You can also store the client file in other directories, and ensure that user **omm** has the read, write, and execute permissions on the specified directory. If the client file already exists in the path, the existing client file will be replaced. ▪ After the file is generated, copy the obtained package to another directory, for example, **/opt/Bigdata/hadoopclient**, as user **omm** or client installation user. When a cluster has many services installed, the cluster client file becomes quite large. Additionally, decompressing this file during installation can consume significant disk space. It is advised to download the client files to a different directory that has ample space, or to promptly remove unnecessary files from the client download directory after installation. Doing so helps avoid exhausting the **/tmp** directory's disk space, which could interrupt the normal operation of the cluster nodes. – If download location is a **Remote Node**: The default directory is **/tmp/FusionInsight-Client-Remote/** on the active OMS node. You can also specify other directories. If there already are client files in the path, they will be overwritten. For a remote node, write permission for the path is required. If you select **Remote Node**, configure the required parameters by referring to **Table 6-2**.	/tmp/FusionInsight-Client/

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 6-2 Parameters for downloading to a remote node

Parameter	Description	Example Value
Host IP Address	IP address of the remote node. The platform type of the remote node must be the same as that of the downloaded client. Otherwise, the client may fail to be installed.	x.x.x.x
Host Port	Host port of the remote node.	22
Username	Username for logging in to the remote node. The user must have the write permission for the path.	xxx
Authentication Method	You can choose one of the following methods: – **Password**: Use the password for login. – **SSH private keys**: Use SSH private keys for login. – **None**: To use this method, passwordless login needs to be enabled.	Password
Password	Enter the password when **Authentication Method** is set to **Password**. Password for logging in to the remote node.	xxx
SSH private keys	This parameter is mandatory when **Authentication Method** is set to **SSH private keys**. Click **Add** and select the SSH private key file to upload.	-
Auto Deployment	Whether to enable auto deployment. This parameter is mandatory when **Select Client Type** is set to **Complete Client**. – If you set this parameter to **yes**, the client is automatically installed and deployed on the current node. – If you set this parameter to **no**, the client will not be automatically installed and deployed. You need to manually install the client after it is downloaded.	Yes
Deployment Path	This parameter is mandatory when **Auto Deployment** is set to **Yes**. If only the configuration file is downloaded, this parameter will not be displayed. The deployment path must be empty if it already exists on the remote node. Otherwise, it will be created automatically. The path also requires operate and write permissions.	/opt/testclient

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

6. If the client is downloaded to the active OMS node, click **Hosts** on FusionInsight Manager to obtain the IP address of the active OMS node.

   If a host name starts with ★, it is the active management node (active OMS node). View and record the value of **Management IP Address** in the row containing the active node.

Step 2 If the client needs to be installed on a node other than the active OMS node in the cluster, copy the client software package to the specified directory on the node where the client is to be installed. **Otherwise, skip this step.**

1. Log in to the active OMS node obtained in **Step 1.6** as user **root**.

   For details about how to log in to the MRS cluster node, see **Logging In to an MRS Cluster Node**.

   After the login is successful, switch to user **omm**.

   ```
   su - omm
   ```

2. Run the following command to copy the software package to the specified directory on the node, for example, **/tmp/clienttemp**:

   ```
   scp -p /tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_Client.tar IP address of the node where the client is to be installed:/tmp/clienttemp
   ```

   If only one component's client is downloaded, the file name format is **FusionInsight_Cluster_1_***<Component name>*_**Client.tar**.

Step 3 Log in to the node where the client is to be installed as the client installation user.

For details about how to log in to the MRS cluster node, see **Logging In to an MRS Cluster Node**.

- You can install the client as user **root** or any OS user. The user must have the operation permission on the client file storage directory and installation directory. For example, the permission on the two directories is **755**. Users **root** and **omm** have the required permission by default. You can enter the required directory and run the **ll** command to view the permission.
- When you install the client as a user other than **omm** or **root**, and the **/var/tmp/patch** directory already exists, you have changed the permission for the directory to **777** and changed the permission for the logs in the directory to **666**. The command for changing the permission is **chmod** *[who] operator [permission] filename*.

Step 4 Decompress the client software package.

1. Go to the path where the client software package is stored.
   - If the client software package is downloaded and needs to be installed on the active OMS node, the path is the value of **Save to Path** configured in **Step 1**, for example, **/tmp/FusionInsight-Client**.
   - If the download location is **Browser**, the path is the directory you configured for the browser to save downloads.
   - If you have performed **Step 2**, the path is the directory you specified, for example, **/tmp/clienttemp**.

   For example, if the client software package is stored in **/tmp/FusionInsight-Client**, run the following command:

   ```
   cd /tmp/FusionInsight-Client
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Decompress the installation package.
   ```
   tar -xvf FusionInsight_Cluster_1_Services_Client.tar
   ```
3. Run the **sha256sum** command to verify the decompressed file.
   ```
   sha256sum -c FusionInsight_Cluster_1_Services_ClientConfig.tar.sha256
   ```
   The command output is as follows:
   ```
   FusionInsight_Cluster_1_Services_ClientConfig.tar: OK
   ```
4. Decompress the installation package.
   ```
   tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar
   ```

Step 5 Go to the directory where the client software package is decompressed and run the following command to install the client to the specified directory:

```
cd FusionInsight_Cluster_1_Services_ClientConfig
```

./install.sh *Client installation directory*

If the specified client installation directory does not exist, the system will automatically create it. If you opt to use an existing directory, it must be empty. The installation directory cannot contain spaces and can contain only letters, digits, and underscores (_).

Example:

```
./install.sh /opt/hadoopclient
```

Wait until the client installation is complete. If the installation is successful, the following information is displayed:

```
...
ALL component client is installed successfully
...
```

- To allow only the current installation user to use the installed client, you can add the **-o** parameter during the installation.

  For example:

  ```
  ./install.sh /opt/hadoopclient -o
  ```
- In MRS 3.3.0 or later, if you do not need to verify clock synchronization during client installation, run the **./install.sh /opt/hadoopclient -u** command to install the client.
- In MRS 3.3.0 or later, if you do not need to verify clock synchronization during client installation and the installed client can be used by the installation user only, add the **-ou** parameter during the installation, that is, run the **./install.sh /opt/hadoopclient -ou** command to install the client.
- If the client needs to be installed on a node outside the cluster and the NTP server mode on the local node is different from that in the cluster, run the **./install.sh /opt/hadoopclient -o chrony** or **./install.sh /opt/hadoopclient -o ntp** command to install the client.

  In the command, **chrony** or **ntp** must be the same as the NTP server mode on the local node.
- If the client node is a node outside the cluster and cannot communicate with the service plane IP address of the active OMS node or cannot access port 20029 of the active OMS node, the client can be successfully installed but cannot be registered with the cluster or displayed in the client list on FusionInsight Manager.
- You need to manually delete the client installation directory when uninstalling a client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Check whether the client is installed.

1. Go to the client installation directory and load environment variables.
   ```
   cd /opt/hadoopclient
   source bigdata_env
   ```
2. Run related commands based on the cluster mode.
   - If Kerberos authentication is not enabled for the cluster, you can directly run commands related to the component client.

     View files in the HDFS root directory:

     ```
     hdfs dfs -ls /
     ```
   - If Kerberos authentication is enabled for the cluster, run the **kinit** command to perform user authentication.

     Example:

     ```
     kinit admin
     ```

     The command output is as follows:

     ```
     Password for xxx@HADOOP.COM: #Enter the password.
     ```

     Run the **klist** command to query and confirm authentication details.

     ```
     klist
     ```

     The command output is as follows:

     ```
     Ticket cache: FILE:/tmp/krb5cc_0
     Default principal: xxx@HADOOP.COM

     Valid starting       Expires              Service principal
     ...
     ```

Step 7 If the previously installed client is no longer available after cluster reinstallation, perform the following operations to deploy the client again:

1. Log in to the node where the client is located as user **root**.
2. Run the following command to view the directory where the client is located: (In the following example, **/opt/hadoopclient** is the directory where the client is located.)
   ```
   ll /opt
   ```

   The command output is as follows:

   ```
   drwxr-x---. 6 root root       4096 Dec 11 19:00 hadoopclient
   drwx------. 2 root root      16384 Nov  6 01:03 lost+found
   drwxr-xr-x. 2 root root       4096 Nov  7 09:49 rh
   ```
3. Run the following command to delete the files in the folder (for example, **/opt/hadoopclient**) where all client programs are located:
   ```
   mv /opt/hadoopclient /tmp/clientbackup
   ```
4. Reinstall the client.

----End

Installing a Client on a Node Outside a Cluster

Step 1 Prepare a Linux ECS for installing the MRS cluster client.

- The recommended ECS OSs and versions are as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 6-3 OS reference list

CPU Architecture	OS	Version
x86 computing	EulerOS	EulerOS 2.5
	SUSE	SUSE Linux Enterprise Server 12 SP4 (SUSE 12.4)
	Red Hat	Red Hat-7.5-x86_64 (Red Hat 7.5)
	CentOS	CentOS 7.6
Kunpeng computing	EulerOS	EulerOS 2.8
	CentOS	CentOS 7.6

- Sufficient disk space (at least 40 GB) must be allocated to the ECS client installation directory.
- The ECS must be in the same VPC and security group as the MRS cluster.
- All ports in the inbound direction of the MRS cluster security group are open to the client node.
- The NTP service has been installed on the ECS OS and is running properly.

  If the NTP service is not installed, run the following command to install it when the **yum** source is configured:

  ```
  yum install ntp -y
  ```
- The ECS must allow users to log in to it using a password (SSH).

Step 2 Configure NTP time synchronization for the node where the client is to be installed to synchronize time with the MRS cluster.

1. Log in to the MRS console.
2. Click an MRS cluster name to go to the cluster details page.
3. Click **Nodes** to expand the Master node group list and view the IP addresses of the **Master1** and **Master2** nodes in the cluster.
4. Log in to the node where the client is to be installed as user **root** and run the following command to edit the NTP configuration file. Add the IP addresses of the **Master1** and **Master2** nodes in the MRS cluster and comment out other **server** addresses.

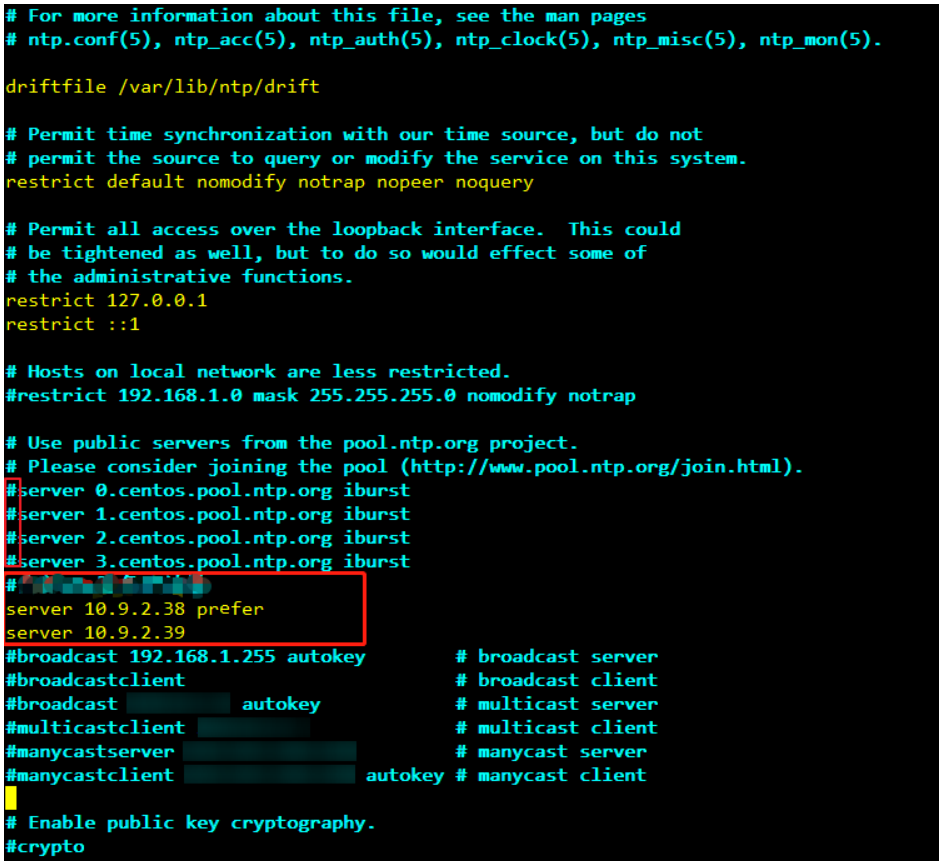
   ```
   vi /etc/ntp.conf
   ```

   Modify the following content:

   ```
   server master1_ip prefer
   server master2_ip
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 6-1 Modifying the NTP configuration file



5. Save the configuration file and run the following command to disable the NTP service:
   ```
   service ntpd stop
   ```
6. Manually synchronize the time.
   ```
   /usr/sbin/ntpdate IP address of the active Master node
   ```
7. Start the NTP service.
   ```
   service ntpd start
   ```
   Or
   ```
   systemctl restart ntpd
   ```
8. Run the **ntpstat** command to check the time synchronization result.
   ```
   ntpstat
   ```
   The command output is as follows:
   ```
   synchronised to NTP server (xxx) at stratum 2
     time correct to within 12 ms
     polling server every 16 s
   ```

Step 3 Obtain the client software package.

Log in to FusionInsight Manager by referring to **Accessing MRS Manager**. Click **Cluster**. On the **Dashboard** page, click **More** and select **Download Client**. In the displayed **Download Client** dialog box, configure parameters and click **OK**.

- The client software package downloaded from the FusionInsight Manager homepage contains the clients of all services (excluding Flume) in the cluster. To download the client of a single service, choose **Cluster** > **Services** > *Service name*, click **More**, and select **Download Client**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 3.3.0 or later, click **Download Client** on the FusionInsight Manager homepage to download the client.

Table 6-4 Client download parameters

Parameter	Description	Example Value
Select Client Type	• **Complete Client**: contains the complete client software package and configuration files. Generally, this option is selected. • **Configuration Files Only**: downloads only client configuration files in the scenario where the administrator modifies the component server configuration on FusionInsight Manager after the complete client is downloaded and installed in an application development task, and developers need to update client configuration files.	Complete Client
Select Platform Type	**The client type must match the architecture of the node where the client is to be installed. Otherwise, the installation fails.** **For clusters of the LTS version, only the client software package whose type is the same as that of FusionInsight Manager can be downloaded.** • **x86_64**: indicates the client software package that can be deployed on an x86 platform. • **aarch64**: indicates the client software package that can be deployed on a Kunpeng server.	x86_64
Save to Path	The path for storing the client software package on the active OMS node • Select **Save to Path**: Customize the path for storing the client software package on the active OMS node. User **omm** must have the read, write, and execute permissions on the path. If the path is not changed, the client file generated is saved in the **/tmp/FusionInsight-Client** directory on the active OMS node in the cluster by default. • Not to select **Save to Path**: The generated client file is automatically downloaded and saved to the local host. Before installing the client, you need to upload the file to a specified directory on the target node.	Select **Save to Path**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Select Download Location	The location for saving the downloaded client file **(only supported by MRS 3.3.1 or later)**. • **Server**: Download the file to the active OMS node of the cluster. • **Browser**: Download the file to the local computer. • **Remote Node:** Download the file to a node other than the active OMS node. If you select a remote location, configure the required parameters by referring to **Table 6-5**.	Server
Save to Path	The path for storing the client software package **(only supported by MRS 3.3.1 or later)**. • If you select **Browser** as the download location, you do not need to configure this parameter. • If the download location is a **Server**: – The default directory is **/tmp/FusionInsight-Client** on the active OMS node. You can also store the client file in other directories, and ensure that user **omm** has the read, write, and execute permissions on the specified directory. If the client file already exists in the path, the existing client file will be replaced. – After the file is generated, copy the obtained package to another directory, for example, **/opt/Bigdata/hadoopclient**, as user **omm** or client installation user. When a cluster has many services installed, the cluster client file becomes quite large. Additionally, decompressing this file during installation can consume significant disk space. It is advised to download the client files to a different directory that has ample space, or to promptly remove unnecessary files from the client download directory after installation. Doing so helps avoid exhausting the **/tmp** directory's disk space, which could interrupt the normal operation of the cluster nodes. • If download location is a **Remote Node**: The default directory is **/tmp/FusionInsight-Client-Remote/** on the active OMS node. You can also specify other directories. If there already are client files in the path, they will be overwritten. For a remote node, write permission for the path is required. If you select **Remote Node**, configure the required parameters by referring to **Table 6-5**.	/tmp/FusionInsight-Client/

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 6-5 Parameters for downloading to a remote node

Parameter	Description	Example Value
Host IP Address	IP address of the remote node. The platform type of the remote node must be the same as that of the downloaded client. Otherwise, the client may fail to be installed.	x.x.x.x
Host Port	Host port of the remote node.	22
Username	Username for logging in to the remote node. The user must have the write permission for the path.	xxx
Authentication Method	You can choose one of the following methods: • **Password**: Use the password for login. • **SSH private keys**: Use SSH private keys for login. • **None**: To use this method, passwordless login needs to be enabled.	Password
Password	Enter the password when **Authentication Method** is set to **Password**. Password for logging in to the remote node.	xxx
SSH private keys	This parameter is mandatory when **Authentication Method** is set to **SSH private keys**. Click **Add** and select the SSH private key file to upload.	-
Auto Deployment	Whether to enable auto deployment. This parameter is mandatory when **Select Client Type** is set to **Complete Client**. • If you set this parameter to **yes**, the client is automatically installed and deployed on the current node. • If you set this parameter to **no**, the client will not be automatically installed and deployed. You need to manually install the client after it is downloaded.	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Deployment Path	This parameter is mandatory when **Auto Deployment** is set to **Yes**. If only the configuration file is downloaded, this parameter will not be displayed. The deployment path must be empty if it already exists on the remote node. Otherwise, it will be created automatically. The path also requires operate and write permissions.	/opt/testclient

Step 4 If the client needs to be installed on a node other than the active OMS node in the cluster, copy the client software package to the specified directory on the node where the client is to be installed. **Otherwise, skip this step.**

1. Log in to the active OMS node obtained in **Step 1.6** as user **root**.

   For details about how to log in to the MRS cluster node, see **Logging In to an MRS Cluster Node**.

   After the login is successful, switch to user **omm**.

   ```
   su - omm
   ```

2. Run the following command to copy the software package to the specified directory on the node, for example, **/tmp/clienttemp**:

   ```
   scp -p /tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_Client.tar IP address of the node where the client is to be installed:/tmp/clienttemp
   ```

   If only one component's client is downloaded, the file name format is **FusionInsight_Cluster_1_***<Component name>*_**Client.tar**.

Step 5 Log in to the node where the client is to be installed as the client installation user.

For details about how to log in to the MRS cluster node, see **Logging In to an MRS Cluster Node**.

- You can install the client as user **root** or any OS user. The user must have the operation permission on the client file storage directory and installation directory. For example, the permission on the two directories is **755**. Users **root** and **omm** have the required permission by default. You can enter the required directory and run the **ll** command to view the permission.
- When you install the client as a user other than **omm** or **root**, and the **/var/tmp/patch** directory already exists, you have changed the permission for the directory to **777** and changed the permission for the logs in the directory to **666**. The command for changing the permission is **chmod** *[who] operator [permission] filename*.

Step 6 Decompress the client software package.

1. Go to the path where the client software package is stored.
   - If the client software package is downloaded and needs to be installed on the active OMS node, the path is the value of **Save to Path** configured in **Step 1**, for example, **/tmp/FusionInsight-Client**.
   - If the download location is **Browser**, the path is the directory you configured for the browser to save downloads.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If you have performed **Step 2**, the path is the directory you specified, for example, **/tmp/clienttemp**.

For example, if the client software package is stored in **/tmp/FusionInsight-Client**, run the following command:

```
cd /tmp/FusionInsight-Client
```

2. Decompress the installation package.
   ```
   tar -xvf FusionInsight_Cluster_1_Services_Client.tar
   ```
3. Run the **sha256sum** command to verify the decompressed file.
   ```
   sha256sum -c FusionInsight_Cluster_1_Services_ClientConfig.tar.sha256
   ```
   The command output is as follows:
   ```
   FusionInsight_Cluster_1_Services_ClientConfig.tar: OK
   ```
4. Decompress the installation package.
   ```
   tar -xvf FusionInsight_Cluster_1_Services_ClientConfig.tar
   ```

Step 7 Check the network connection of the client.

To install the client on a node, map the host names and IP addresses of all nodes in the cluster in the **/etc/hosts** file. Here are the steps to import the domain name mapping of the cluster to the **hosts** file:

1. Switch to user **root** or a user who has the permission to modify the **hosts** file.
   ```
   su - root
   ```
2. Go to the directory where the client package is decompressed.
   ```
   cd /tmp/clienttemp/FusionInsight_Cluster_1_Services_ClientConfig
   ```
3. Import the domain name mapping to the **hosts** file.
   ```
   cat realm.ini >> /etc/hosts
   ```

NOTE

- If the host where the client is installed is not a node in the cluster, configure network connections for the client to prevent errors when you run commands on the client.
- If Spark tasks are executed in yarn-client mode, add the **spark.driver.host** parameter to the file *Client installation directory*/**Spark/spark/conf/spark-defaults.conf** and set the parameter to the client IP address.
- If the yarn-client mode is used, you need to configure the mapping between the IP address and host name of the client in the **hosts** file on the active and standby YARN nodes (ResourceManager nodes in the cluster) to make sure that the Spark Web UI is properly displayed.

Step 8 Go to the directory where the client software package is decompressed and run the following command to install the client to the specified directory:

```
cd FusionInsight_Cluster_1_Services_ClientConfig
```

./install.sh *Client installation directory*

If the specified client installation directory does not exist, the system will automatically create it. If you opt to use an existing directory, it must be empty. The installation directory cannot contain spaces and can contain only letters, digits, and underscores (_).

Example:

```
./install.sh /opt/hadoopclient
```

Wait until the client installation is complete. If the installation is successful, the following information is displayed:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
...
ALL component client is installed successfully
...
```

- To allow only the current installation user to use the installed client, you can add the **-o** parameter during the installation.

  For example:

  ```
  ./install.sh /opt/hadoopclient -o
  ```

- In MRS 3.3.0 or later, if you do not need to verify clock synchronization during client installation, run the **./install.sh /opt/hadoopclient -u** command to install the client.
- In MRS 3.3.0 or later, if you do not need to verify clock synchronization during client installation and the installed client can be used by the installation user only, add the **-ou** parameter during the installation, that is, run the **./install.sh /opt/hadoopclient -ou** command to install the client.
- If the client needs to be installed on a node outside the cluster and the NTP server mode on the local node is different from that in the cluster, run the **./install.sh /opt/hadoopclient -o chrony** or **./install.sh /opt/hadoopclient -o ntp** command to install the client.

  In the command, **chrony** or **ntp** must be the same as the NTP server mode on the local node.
- If the client node is a node outside the cluster and cannot communicate with the service plane IP address of the active OMS node or cannot access port 20029 of the active OMS node, the client can be successfully installed but cannot be registered with the cluster or displayed in the client list on FusionInsight Manager.
- You need to manually delete the client installation directory when uninstalling a client.

Step 9 Check whether the client is installed.

1. Go to the client installation directory and load environment variables.

   ```
   cd /opt/hadoopclient
   source bigdata_env
   ```

2. Run related commands based on the cluster mode.
   - If Kerberos authentication is not enabled for the cluster, you can directly run commands related to the component client.

     View files in the HDFS root directory:

     ```
     hdfs dfs -ls /
     ```

   - If Kerberos authentication is enabled for the cluster, run the **kinit** command to perform user authentication.

     Example:

     ```
     kinit admin
     ```

     The command output is as follows:

     ```
     Password for xxx@HADOOP.COM: #Enter the password.
     ```

     Run the **klist** command to query and confirm authentication details.

     ```
     klist
     ```

     The command output is as follows:

     ```
     Ticket cache: FILE:/tmp/krb5cc_0
     Default principal: xxx@HADOOP.COM
     ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
Valid starting     Expires          Service principal
...
```

----End

Helpful Links

- For details about how to install a client in a cluster of MRS 2.x or earlier, see **Installing a Client (MRS 2.x or Earlier)**.
- After downloading and installing the client, you can view information about the installed (registered) client on FusionInsight Manager. For details, see **Viewing the Installed MRS Cluster Client**.
- If you have modified service parameters and restarted the service, you need to download and install the installed client again or use the configuration file to update the client. For details about how to update the client, see **Updating the MRS Cluster Client After the Server Configuration Expires**.
- If multiple clients need to be upgraded after cluster upgrade or scaling, you can upgrade the clients in batches by referring to **Batch Upgrading MRS Cluster Clients**.

6.2 Installing a Client (MRS 2.x or Earlier)

You need to install a cluster client to connect to the component server in the cluster and perform tasks such as component connection and job submission. You can install the cluster client on a node within the cluster or on a node outside of it.

To ensure that certain features function properly, it is important to reinstall the client after modifying the server configuration of a component in the cluster. This ensures that the client version matches the server version.

This section describes how to install the cluster client of MRS 2.*x* or earlier. For details about MRS 3.*x*, see **Installing a Client (MRS 3.x)**.

After an MRS cluster earlier than 3.x is created, the client is installed on the active Master node by default. The installation directory is **/opt/client**.

Installing a Client on a Node Inside a Cluster

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Services** > **Download Client**, and download the client installation package to the active OMS node.

Step 3 On the MRS console, view the IP address on the **Nodes** page of the specified cluster.

Record the IP address of the node where the client is to be installed and the IP address of the active Master node (active OMS node).

Step 4 Log in to the active OMS node as user **root** and run the following command to switch to user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
sudo su - omm
```

Step 5 Copy the client installation package to the specified node.

```
scp -p /tmp/MRS-client/MRS_Services_Client.tar IP address of the node where the client is to be installed:/opt/client_tmp
```

Step 6 Log in to the node where the client is to be installed as the **root** user.

Step 7 Install the client.

```
cd /opt/client_tmp
tar -xvf MRS_Services_Client.tar
tar -xvf MRS_Services_ClientConfig.tar
cd MRS_Services_ClientConfig
./install.sh Client installation directory
```

Example:

```
./install.sh /opt/hadoopclient
```

Wait until the client is installed. If the following information is displayed, the client is installed successfully:

```
Components client installation is complete.
```

Client installation directory:

- **If the specified client installation directory does not exist, the system will automatically create it. If you opt to use an existing directory, it must be empty. The installation directory cannot contain spaces and can contain only letters, digits, and underscores (_).**
- You need to manually delete the client installation directory when uninstalling a client.

Step 8 Check whether the client is installed.

1. Go to the client installation directory and load environment variables.
   ```
   cd /opt/hadoopclient
   source bigdata_env
   ```
2. Run related commands based on the cluster mode.
   - If Kerberos authentication is not enabled for the cluster, you can directly run commands related to the component client.

     View files in the HDFS root directory:

     ```
     hdfs dfs -ls /
     ```
   - If Kerberos authentication is enabled for the cluster, run the **kinit** command to perform user authentication.

     Example:

     ```
     kinit admin
     ```

     The command output is as follows:

     ```
     Password for xxx@HADOOP.COM: #Enter the password.
     ```

     Run the **klist** command to query and confirm authentication details.

     ```
     klist
     ```

     The command output is as follows:

     ```
     Ticket cache: FILE:/tmp/krb5cc_0
     Default principal: xxx@HADOOP.COM

     Valid starting     Expires            Service principal
     ...
     ```

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Installing a Client on a Node Outside a Cluster

Step 1 Prepare a Linux ECS for installing the MRS cluster client.

- The recommended ECS OSs and versions are as follows:

  Table 6-6 OS reference list

CPU Architecture	OS	Supported Version
x86 computing	EulerOS	– Available: EulerOS 2.2 – Available: EulerOS 2.3 – Available: EulerOS 2.5

- The CPU architecture of the ECS must be the same as that of the MRS cluster node.
- Sufficient disk space (at least 40 GB) must be allocated to the ECS client installation directory.
- The ECS must be in the same VPC and security group as the MRS cluster.
- All ports in the inbound direction of the MRS cluster security group are open to the client node.
- The NTP service has been installed on the ECS OS and is running properly.

  If the NTP service is not installed, run the **yum install ntp -y** command to install it when the **yum** source is configured.
- The ECS must allow users to log in to it using a password (SSH).

Step 2 Log in to Manager of the cluster by referring to **Accessing MRS Manager** and choose **Services** > **Download Client**.

- In **Client Type**, select **All client files**.
- In **Download to**, select **Remote host**.
- Set **Host IP Address** to the IP address of the node where the client is to be installed, **Host Port** to **22**, and **Save Path** to **/tmp**.

  If the default SSH login port of the node has been changed, set **Host Port** to the new port.
- Set **Login User** to **root**.

  If other users are used, ensure that the users have read, write, and execute permission on the save path.
- Select **Password** or **SSH Private Key** for **Login Mode**.
  - **Password**: Enter the password of user **root** set during cluster creation.
  - **SSH Private Key**: Select and upload the key file used for creating the cluster.

Step 3 Click **OK** to generate a client file.

After the file download is confirmed, obtain the client software package from the specified path on the node where the client will be installed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the client download fails, check the username, password, and security group configuration of the remote host. Ensure that the username and password are accurate and that the inbound rule for the SSH port has been added to the remote host's security group. Once confirmed, retry the client download.

Generating a client will occupy a large number of disk I/Os. You are advised not to download a client when the cluster is being installed, started, and patched, or in other unstable states.

Step 4 Configure NTP time synchronization for the node where the client is to be installed to synchronize time with the MRS cluster.

1. Log in to the MRS console.
2. Click an MRS cluster name to go to the cluster details page.
3. Click **Nodes** to expand the Master node group list and view the IP addresses of the **Master1** and **Master2** nodes in the cluster.
4. Log in to the node where the client is to be installed as user **root** and run the following command to edit the NTP configuration file. Add the IP addresses of the **Master1** and **Master2** nodes in the MRS cluster and comment out other **server** addresses.

   ```
   vi /etc/ntp.conf
   ```

   Modify the following content:

   ```
   server master1_ip prefer
   server master2_ip
   ```

   Figure 6-2 Modifying the NTP configuration file

   ```
   # For more information about this file, see the man pages
   # ntp.conf(5), ntp_acc(5), ntp_auth(5), ntp_clock(5), ntp_misc(5), ntp_mon(5).

   driftfile /var/lib/ntp/drift

   # Permit time synchronization with our time source, but do not
   # permit the source to query or modify the service on this system.
   restrict default nomodify notrap nopeer noquery

   # Permit all access over the loopback interface.  This could
   # be tightened as well, but to do so would effect some of
   # the administrative functions.
   restrict 127.0.0.1
   restrict ::1

   # Hosts on local network are less restricted.
   #restrict 192.168.1.0 mask 255.255.255.0 nomodify notrap

   # Use public servers from the pool.ntp.org project.
   # Please consider joining the pool (http://www.pool.ntp.org/join.html).
   #server 0.centos.pool.ntp.org iburst
   #server 1.centos.pool.ntp.org iburst
   #server 2.centos.pool.ntp.org iburst
   #server 3.centos.pool.ntp.org iburst
   #
   server 10.9.2.38 prefer
   server 10.9.2.39
   #broadcast 192.168.1.255 autokey        # broadcast server
   #broadcastclient                        # broadcast client
   #broadcast                  autokey     # multicast server
   #multicastclient                        # multicast client
   #manycastserver                         # manycast server
   #manycastclient                 autokey # manycast client

   # Enable public key cryptography.
   #crypto
   ```

5. Save the configuration file and run the following command to disable the NTP service:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
service ntpd stop
```

6. Manually synchronize the time.
   ```
   /usr/sbin/ntpdate IP address of the active Master node
   ```
7. Start the NTP service.
   ```
   service ntpd start
   ```
   Or
   ```
   systemctl restart ntpd
   ```
8. Run the **ntpstat** command to check the time synchronization result.
   ```
   synchronised to NTP server (xxx) at stratum 2
     time correct to within 12 ms
     polling server every 16 s
   ```

Step 5 Copy the downloaded client software package to the **/opt** directory.

```
cp /tmp/MRS_Services_Client.tar /opt
```

Step 6 Decompress the software package.

```
cd /opt
tar -xvf MRS_Services_Client.tar
```

Step 7 Verify the file.

```
sha256sum -c MRS_Services_ClientConfig.tar.sha256
```

The command output is as follows:

```
MRS_Services_ClientConfig.tar: OK
```

Step 8 Decompress the software package.

```
tar -xvf MRS_Services_ClientConfig.tar
```

Step 9 Install the cluster client.

```
sh /opt/MRS_Services_ClientConfig/install.sh Client installation directory
```

Example:

```
sh /opt/MRS_Services_ClientConfig/install.sh /opt/hadoopclient
```

If the following information is displayed, the client has been successfully installed:

```
Components client installation is complete.
```

Client installation directory:

- **If the specified client installation directory does not exist, the system will automatically create it. If you opt to use an existing directory, it must be empty. The installation directory cannot contain spaces and can contain only letters, digits, and underscores (_).**
- You need to manually delete the client installation directory when uninstalling a client.

Step 10 Check whether the client is installed.

1. Go to the client installation directory and load environment variables.
   ```
   cd /opt/hadoopclient
   source bigdata_env
   ```
2. Run related commands based on the cluster mode.
   - If Kerberos authentication is not enabled for the cluster, you can directly run commands related to the component client.

     View files in the HDFS root directory:

     ```
     hdfs dfs -ls /
     ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If Kerberos authentication is enabled for the cluster, run the **kinit** command to perform user authentication.

  Example:

  ```
  kinit admin
  ```

  The command output is as follows:

  ```
  Password for xxx@HADOOP.COM: #Enter the password.
  ```

  Run the **klist** command to query and confirm authentication details.

  ```
  klist
  ```

  The command output is as follows:

  ```
  Ticket cache: FILE:/tmp/krb5cc_0
  Default principal: xxx@HADOOP.COM

  Valid starting     Expires            Service principal
  ...
  ```

----End

Helpful Links

- For details about how to install a client in a cluster of MRS 3.x, see **Installing a Client (MRS 3.x)**.
- If you have modified service parameters and restarted the service, you need to download and install the installed client again or use the configuration file to update the client. For details about how to update the client, see **Updating the MRS Cluster Client After the Server Configuration Expires**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

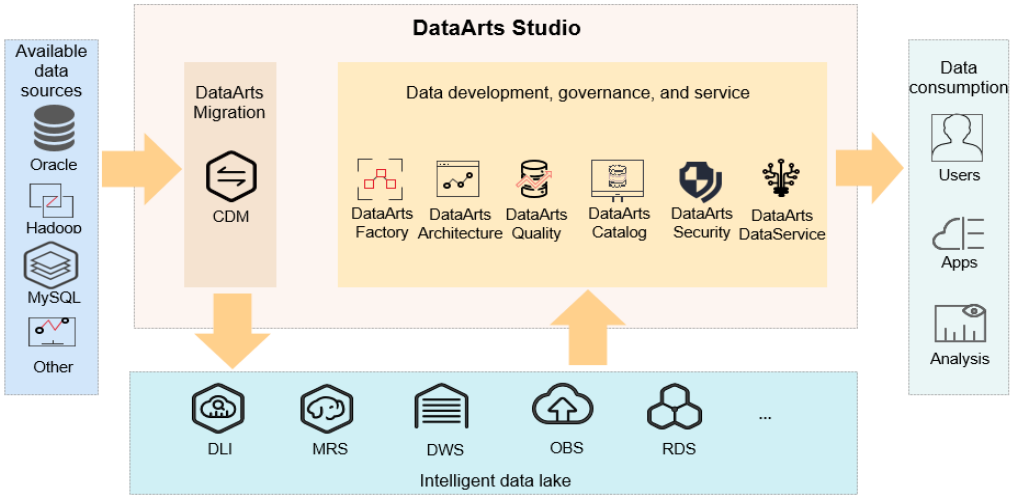
7 Submitting an MRS Job

7.1 MRS Job Introduction

After an MRS cluster is created, you can develop upper-layer applications or scripts based on service requirements to process service data. You can develop job programs or scripts based on big data component capabilities and actual service requirements, upload the programs or scripts together with service data to the cloud platform, and run the programs or scripts for specific services.

- MRS clusters enable online job management. You can create and run jobs on the MRS console. The MRS service supports MapReduce, Spark, Hive SQL, and Spark SQL jobs. You can also use the cluster client to submit jobs to a big data component server.
- DataArts Studio provides a one-stop big data collaboration development environment and fully managed big data scheduling capabilities. With DataArts Studio, you can develop and debug MRS Hive SQL/Spark SQL scripts online, develop MRS application jobs by performing drag-and-drop operations, and migrate and integrate data between MRS and other heterogeneous data sources.

Figure 7-1 DataArts Studio architecture

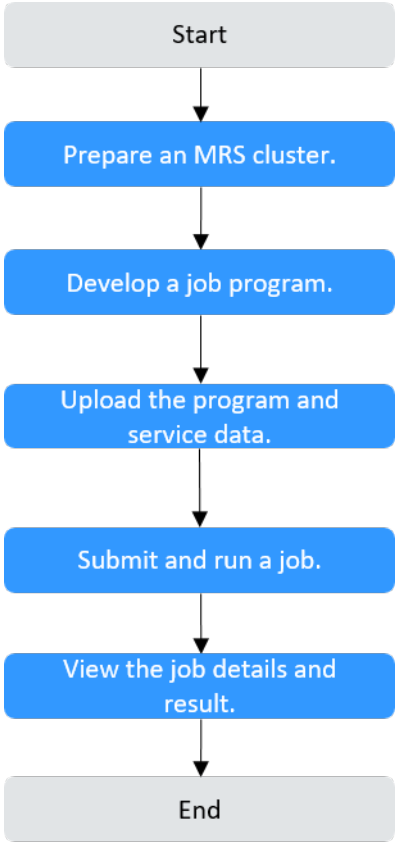


 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

DataArts Studio and MRS can be used together to build a full-link big data solution encompassing data ingestion, processing, and analytics.

- During enterprise data lake building, DataArts Studio can use the data integration module to collect data from multi-source heterogeneous systems, use the data development module to compile Hive, Spark, and Flink scripts and submit the scripts to the MRS cluster for distributed computing, and use MRS HDFS or HBase to store massive amounts of raw data. Meanwhile, the data governance function provided by DataArts Studio can also be used to manage metadata, monitors data quality, and traces data lineage.
- During real-time data processing, DataArts Studio can orchestrate streaming jobs, connect to MRS Kafka to obtain real-time data streams, and perform real-time data cleaning and aggregation using the MRS Spark Streaming or Flink cluster. The results are stored in the MRS Hive data warehouse. The report module of DataArts Studio then performs dynamic and visualized analysis based on the results.
- During offline batch processing, DataArts Studio can schedule MRS Hadoop MapReduce jobs to complete ETL processing of large-scale data, such as transaction reconciliation in the financial industry and device log analysis in the manufacturing industry. In addition, you can use the intelligent data modeling function of DataArts Studio to build business models for data stored in MRS, achieving integrated operation of data value mining and compliance governance.

Figure 7-2 MRS Job Submission Process



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Data processed by MRS jobs is usually stored in OBS or HDFS. Before creating a job, you need to upload the data to be analyzed to an OBS file system or the HDFS in the MRS cluster. The MRS console also allows you to quickly import data from OBS to HDFS. HDFS and OBS can store compressed data in **.bz2** and **.gz** formats.

MRS clusters support the following job types:

- MapReduce can quickly process large-scale data in parallel. It is a distributed data processing model and execution environment. MRS supports the submission of MapReduce JAR programs.
- Spark is a distributed in-memory computing framework. MRS supports SparkSubmit, Spark Script, and Spark SQL jobs.
  - SparkSubmit: You can submit Spark JAR and Spark Python programs, execute the Spark Application, and compute and process user data.
  - SparkScript: You can submit SparkScript scripts and batch execute Spark SQL statements.
  - Spark SQL: You can use Spark SQL statements (similar to SQL statements) to query and analyze user data in real time.
- Hive is an open-source data warehouse based on Hadoop. MRS allows you to submit HiveScript scripts and directly execute Hive SQL statements.
- Flink is a distributed big data processing engine that can perform stateful computations over both unbounded and bounded data streams.
- Hadoop Streaming works similarly to a standard Hadoop job, where you can define the input and output HDFS paths, as well as the mapper and reducer executable programs.

User Permissions Required for Running Jobs on the MRS Console

Before submitting a job on the MRS console, you need to synchronize IAM users to the user system of the MRS cluster. Whether a synchronized IAM user has the permission to submit jobs depends on the IAM policy associated with the user during IAM synchronization. For details about the job submission policy, see **Table 8-62** in **Synchronizing IAM Users to MRS**.

When a user submits a job that involves the resource usage of a specific component, such as accessing HDFS directories and Hive tables, the FusionInsight Manager administrator (for example, user **admin**) must grant the required permissions to the user.

Step 1 Log in to FusionInsight Manager of the cluster as user **admin**. For details, see **Accessing MRS Manager**.

Step 2 Add the role of the component whose permission is required by the user. For details, see **Managing MRS Cluster Roles**.

Step 3 Change the user group to which the user who submits the job belongs and add the new component role to the user group. For details, see **Managing MRS Cluster User Groups**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the component role associated with the user group to which the user belongs is modified, it takes at least 5 minutes for the role permissions to take effect.

----End

7.2 Uploading Application Data to an MRS Cluster

Data sources processed by an MRS cluster are usually from the OBS or HDFS file system in the cluster. You can browse, manage, and use OBS data on the MRS console or using the OBS client. Alternatively, you can import OBS data to the HDFS file system in the cluster for processing.

Generally, you are advised to store service data in the OBS file system. When data in the OBS file system is being migrated to the HDFS file system, file transfer speeds generally decrease as file size increases. Therefore, it is recommended that you import data from OBS to the HDFS file system only when there is a small amount of data.

Uploading Local Data to OBS

Step 1 Log in to the OBS console.

Step 2 Choose **Parallel File Systems** > **Create Parallel File System**.

Step 3 Enter the file system name, for example, **mrs-demotest**.

Step 4 Click **Create Now**.

Step 5 In the parallel file system list on the OBS console, click the file system name to go to its details page.

Step 6 In the navigation pane on the left, choose **Files**. Create different folders based on service requirements and upload data to the folders.

For example, create the following folders:

- **program**: Upload the program package for running to this folder.
- **input**: Upload the input data to this folder.
- **output**: Job running results will be written to this folder.

----End

Uploading Local Data on a Cluster Node to HDFS

Use the HDFS client to upload the local program or service data on a node in the MRS cluster to HDFS.

Step 1 Install an MRS cluster client.

For details, see **Installing an MRS Cluster Client**.

The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Log in to the node where the client is located as the MRS cluster client installation user.

For details, see **Logging In to an MRS Cluster Node**.

Step 3 Run the following command to go to the client installation directory and load the environment variables:

```
source /opt/Bigdata/client/bigdata_env
```

Step 4 For a cluster with Kerberos authentication enabled (security mode), run the following command to authenticate the user (the user who logs in for the first time needs to change the password). Skip this step for a cluster in normal mode.

Ensure that you have created a user with the HDFS operation permissions in the MRS cluster. For details, see **Creating an MRS Cluster User**.

```
kinit Username of the service user who has the HDFS operation permissions
```

Step 5 Run the following command to create a folder in HDFS:

```
hdfs dfs -mkdir /tmp/mydir
```

Step 6 Upload the local file on the node to the target folder in HDFS.

For example, run the following command to upload the **/tmp/testAccess.txt** file on the node to the **/tmp/mydir** folder in HDFS:

```
hdfs dfs -put /tmp/testAccess.txt /tmp/mydir
```

----End

Importing Data from OBS to HDFS

After an MRS cluster is created, import files from an OBS parallel file system to the HDFS in the cluster using the management console.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to go to the cluster information page.

If the MRS cluster is a security-mode cluster with Kerberos authentication enabled, synchronize IAM users first. (On the **Dashboard** tab page of the cluster details page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.) After synchronizing IAM users for the first time, you are advised to wait for 5 minutes before performing subsequent operations.

Step 3 Click **Files** to go to the file management page.

Step 4 Select **HDFS File List**.

Step 5 Go to the directory where the data to be imported is stored.

Click **Create** to create a folder directory or select an existing folder in HDFS.

Step 6 Click **Import Data** and configure the HDFS and OBS paths correctly.

When configuring the OBS or HDFS path, click **Browse**, select a file directory, and click **OK**.

- OBS path description:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - The path must start with **obs://**.
  - Files or programs encrypted by KMS cannot be imported.
  - An empty folder cannot be imported.
  - The directory and file name can contain letters, digits, hyphens (-), and underscores (_), but cannot contain the following special characters: ;| &>,<'$*?\
  - The directory and file name cannot start or end with a space, but can contain spaces between them.
  - The OBS full path contains a maximum of 255 characters.
- HDFS path description:
  - The directory and file name can contain letters, digits, hyphens (-), and underscores (_), but cannot contain the following special characters: ;| &>,<'$*?\:
  - The directory and file name cannot start or end with a space, but can contain spaces between them.
  - The HDFS full path contains a maximum of 255 characters.

Step 7 Click **OK**.

You can view the file upload progress on the **File Operation Records** page. The system generates a DistCp job for processing the data import operation. You can also view the job execution status on the **Job Management** page.

----End

Exporting Data from HDFS to OBS

After an MRS cluster is created, export specified files from the HDFS in the cluster to the OBS using the management console.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to enter the cluster details page.

If the MRS cluster is a security-mode cluster with Kerberos authentication enabled, synchronize IAM users first. (On the **Dashboard** tab page of the cluster details page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.) After synchronizing IAM users for the first time, you are advised to wait for 5 minutes before performing subsequent operations.

Step 3 Click **Files** to go to the file management page.

Step 4 Select **HDFS File List**.

Step 5 Go to the data storage directory.

Step 6 Click **Export Data** and configure the OBS and HDFS paths. When configuring the OBS or HDFS path, click **Browse**, select a file directory, and click **OK**.

When a folder is exported to OBS, a label file named **folder name_$folder$** is added under the OBS path. Ensure that the exported folder is not empty. If the exported folder is empty, OBS cannot display the folder and only generates a file named **folder name_$folder$**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Click **OK**.

You can view the file upload progress on the **File Operation Records** page. The system generates a DistCp job for processing the data import operation. You can also view the job execution status on the **Job Management** page.

----End

7.3 Submitting a Job in an MRS Cluster

7.3.1 Running a MapReduce Job

MapReduce is a programming model for large-scale data processing. It divides the complex data processing task into two main phases: "map" and "reduce". The map phase splits a large amount of data into small chunks for parallel processing. Each node independently executes the user-defined map function and produces intermediate key/value pairs. The reduce phase groups all values with the same key, uses the reduce function to aggregates and summarizes the values, and produces a final set of key/value pairs. This distributed computing framework efficiently processes PB-scale data while overcoming the performance limitations of traditional single-node processing through fault tolerance and automatic task scheduling. It is widely used in scenarios such as search engine indexing, log analysis, and data statistics.

MRS allows you to submit and run your own programs, and get the results. This section will show you how to submit a MapReduce job in an MRS cluster.

You can create a job online and submit it for running on the MRS console, or submit a job in CLI mode on the MRS cluster client.

Prerequisites

- You have uploaded the program packages and data files required by jobs to OBS or HDFS.
- If the job program needs to read and analyze data in the OBS file system, you need to configure storage-compute decoupling for the MRS cluster. For details, see **Configuring Storage-Compute Decoupling for an MRS Cluster**.

Notes and Constraints

- When the policy of the user group to which an IAM user belongs changes from **MRS ReadOnlyAccess** to **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator**, or vice versa, wait for five minutes after user synchronization for the System Security Services Daemon (SSSD) cache of the cluster node to refresh. Submit a job on the MRS console after the new policy takes effect. Otherwise, the job submission may fail.
- If the IAM username contains spaces (for example, **admin 01**), you cannot create jobs on the MRS console.

Submitting a Job

You can create and run jobs online using the management console or submit jobs by running commands on the cluster client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Submitting a Job on the Console

Step 1 Prepare the application and data.

This section uses the Hadoop word count application as an example. You can obtain the sample program from the MRS cluster client (*Client installation directory*/**HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-***XXX***.jar**) and upload the program to a specified directory in HDFS or OBS. For details, see **Uploading Application Data to an MRS Cluster**.

To run the application, you need to specify the following parameters:

- Program class name: It is specified by a function in your program. In this application, the class name is **wordcount**.
- Input file path: Path of the data file to be analyzed. The file must be uploaded to the HDFS or OBS file system in advance.

  For example, upload data file **data1.txt**. The file content is as follows:

  ```
  This is a test job.
  MRS supports multiple job submission modes.
  ```

- Output file path: Path of the result file after the application counts words. **Set this parameter to a directory that does not exist**. The directory will be automatically generated after you run the application.

Step 2 Log in to the MRS console.

Step 3 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 4 On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.

Perform this step only when Kerberos authentication is enabled for the cluster.

After IAM user synchronization, wait for five minutes before submitting a job. For details about IAM user synchronization, see Synchronizing IAM Users to MRS.

Step 5 Click **Job Management**. On the displayed job list page, click **Create**.

Step 6 In **Type**, select **MapReduce**. Configure other job information.

Table 7-1 Job parameters

Parameter	Description	Example
Name	Job name. It can contain 1 to 128 characters. Only letters, digits, hyphens (-), and underscores (_) are allowed.	mapreduce_job

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
Program Path	Path of the program package to be executed. You can enter the path or click **HDFS** or **OBS** to select a file. ● The path can contain a maximum of 1,023 characters. It cannot contain special characters (;|&>,<'$) and cannot be left blank or all spaces. ● The OBS program path starts with **obs://**. The HDFS program path starts with **hdfs://hacluster**, for example, **hdfs://hacluster/user/*XXX*.jar**. ● The MapReduce job execution program must end with **.jar**.	obs://mrs-demotest/program/hadoop-mapreduce-examples-*XXX*.jar
Runtime Parameters	(Optional) Key parameters for program execution. Use spaces to separate multiple parameters. Configuration format for the example program in this section: *Program class name Data input path Data output path* ● Program class name: It is specified by a function in your program. MRS is responsible for transferring parameters only. ● Data input path: Click **HDFS** or **OBS** to select a path or manually enter a correct path. ● Data output path: output path of the data processing result. Enter a directory that does not exist. The parameter can contain a maximum of 150,000 characters. It cannot contain special characters ;|&><'$, but can be left blank. **CAUTION** When entering a parameter containing sensitive information (for example, login password), you can add an at sign (@) before the parameter name to encrypt the parameter value. This prevents the sensitive information from being persisted in plaintext. When you view job information on the MRS console, the sensitive information is displayed as *****. Example: **username=***testuser* **@password=***User password*	wordcount obs://mrs-demotest/input/data1.txt obs://mrs-demotest/output/demo1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
Service Parameter	(Optional) Service parameters for the job. To modify the current job, change this parameter. For permanent changes to the entire cluster, refer to **Modifying the Configuration Parameters of an MRS Cluster Component** and modify the cluster component parameters accordingly. For example, if decoupled storage and compute is not configured for the MRS cluster and jobs need to access OBS using AK/SK, you can add the following service parameters: • **fs.obs.access.key**: key ID for accessing OBS. • **fs.obs.secret.key**: key corresponding to the key ID for accessing OBS.	-
Command Reference	Commands submitted to the background for execution when a job is submitted.	N/A

Step 7 Confirm job configuration information and click **OK**.

Step 8 After the job is submitted, you can view the job running status and execution result in the job list. After the job status changes to **Completed**, you can view the analysis result of related programs.

In this example, you can view the data statistics in the specified OBS output directory.

During job execution, you can click **View Log** or choose **More** > **View Details** to view program execution details. If the job execution is abnormal or fails, you can locate the fault based on the error information.

A created job cannot be modified. If you need to execute the job again, you can click **Clone** to quickly copy the created job and adjust required parameters.

----End

Submitting a Job Using the Cluster Client

Step 1 Prepare the application and data.

This section uses the Hadoop word count application as an example. You can obtain the sample program from the MRS cluster client (*Client installation directory*/**HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-***XXX***.jar**) and upload the program to a specified directory in HDFS or OBS. For details, see **Uploading Application Data to an MRS Cluster**.

To run the application, you need to specify the following parameters:

- Program class name: It is specified by a function in your program. In this application, the class name is **wordcount**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Input file path: Path of the data file to be analyzed. The file must be uploaded to the HDFS or OBS file system in advance.

  For example, upload data file **data1.txt**. The file content is as follows:

  ```
  This is a test job.
  MRS supports multiple job submission modes.
  ```

- Output file path: Path of the result file after the application counts words. **Set this parameter to a directory that does not exist**. The directory will be automatically generated after you run the application.

Step 2 If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.

Step 3 Install an MRS cluster client.

For details, see **Installing an MRS Cluster Client**.

The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

Step 4 Log in to the node where the client is located as the MRS cluster client installation user.

For details, see **Logging In to an MRS Cluster Node**.

Step 5 Run the following command to go to the client installation directory:

```
cd /opt/Bigdata/client
```

Run the following command to load the environment variables:

```
source bigdata_env
```

If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, you do not need to run the **kinit** command.

```
kinit testuser
```

Step 6 Run a command to submit the word count job.

You can run the **yarn** or **hadoop jar** command to submit a MapReduce job.

yarn commands are primarily used for resource management and job scheduling. These commands allow you to manage resources in a YARN cluster and control the lifecycle of jobs. You can run these commands to submit MapReduce and Spark jobs and query and monitor jobs.

hadoop jar commands are mainly used to submit Java-based MapReduce jobs. These commands enable you to start and run MapReduce jobs, use the main class within the JAR file as the entry point of the jobs, and pass parameters such as the input and output paths to the jobs.

To use the **hadoop jar** command to submit a sample program job, the command format is as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
hadoop jar Application wordcount Input file path Output file path
```

For example, run the following command to use the sample program to count the number of words in the **/tmp/data/data1.txt** file in HDFS and export the result to the **/tmp/output/demo** directory in the HDFS:

```
hadoop jar HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-*.jar wordcount /tmp/data/data1.txt /tmp/output/demo
```

If decoupled storage and compute is not configured for the MRS cluster, jobs need to access OBS using AK/SK. Example command is as follows:

```
hadoop jar HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-*.jar wordcount -Dfs.obs.access.key=Access key ID for accessing OBS -Dfs.obs.secret.key=Secret access key corresponding to the access key ID for accessing OBS "obs://mrs-demotest/input/data1.txt" "obs://mrs-demotest/output/demo"
```

- Commands carrying authentication passwords pose security risks. Disable historical command recording before running such commands to prevent information leakage.
- To obtain the AK and SK, log in to the OBS console and choose **My Credentials** > **Access Keys** from the username drop-down list in the upper right corner of the page.

Step 7 After the job is successfully submitted and executed, run the following command to view the data statistics result in the specified HDFS output directory:

```
hdfs dfs -ls /tmp/output
```

Step 8 Log in to FusionInsight Manager as user **testuser**, choose **Cluster** > **Services** > **Yarn**, and click the hyperlink on the right of **ResourceManager Web UI** to access the YARN Web UI. Click the application ID of the job to view the job running information and related logs.

----End

7.3.2 Running a SparkSubmit Job

Spark is an open-source parallel data processing framework. It helps you easily and quickly develop big data applications and perform offline processing, stream processing, and interactive analysis on data. spark-submit is a utility provided by Spark for deploying and running Spark applications in different clusters. You can use this tool to submit compiled Spark jobs to YARN or execute them locally. It achieves this by packaging application code into independent JAR or WAR files and then leveraging the resource management capabilities of the distributed computing framework for efficient Spark job submitting and execution.

MRS allows you to submit and run your own programs, and get the results. This section will show you how to submit a SparkSubmit job in an MRS cluster.

You can create a job online and submit it for running on the MRS console, or submit a job in CLI mode on the MRS cluster client.

Prerequisites

- You have uploaded the program packages and data files required for running jobs to OBS or HDFS.
- If the job program needs to read and analyze data in the OBS file system, you need to configure storage-compute decoupling for the MRS cluster. For details, see **Configuring Storage-Compute Decoupling for an MRS Cluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Notes and Constraints

- When the policy of the user group to which an IAM user belongs changes from **MRS ReadOnlyAccess** to **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator**, or vice versa, wait for five minutes after user synchronization for the System Security Services Daemon (SSSD) cache of the cluster node to refresh. Submit a job on the MRS console after the new policy takes effect. Otherwise, the job submission may fail.
- If the IAM username contains spaces (for example, **admin 01**), you cannot create jobs on the MRS console.

Submitting a Job

You can create and run jobs online using the management console or submit jobs by running commands on the cluster client.

Submitting a Job on the Console

Step 1 Prepare the application and data.

This section uses the spark-examples sample program provided by the cluster as an example. The sample program provides multiple functions. You can obtain the sample program from the MRS cluster client (*Client installation directory*/**Spark/spark/examples/jars/spark-examples-***XXX***.jar**). In some versions, the name of the **Spark2x** folder in the cluster is **Spark**. Replace it with the actual name. Upload the obtained sample program to a specified directory in HDFS or OBS. For details, see **Uploading Application Data to an MRS Cluster**.

Program class name: It is specified by a function in your program. In this application, the class name is **SparkPi** for calculating the value of Pi (π).

Step 2 Log in to the MRS console.

Step 3 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 4 On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.

Perform this step only when Kerberos authentication is enabled for the cluster.

After IAM user synchronization, wait for five minutes before submitting a job. For details about IAM user synchronization, see Synchronizing IAM Users to MRS.

Step 5 Click **Job Management**. On the displayed job list page, click **Create**.

Step 6 In **Type**, select **SparkSubmit**. Configure other job information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 7-2 Job parameters

Parameter	Description	Example
Name	Job name. It can contain 1 to 128 characters. Only letters, digits, hyphens (-), and underscores (_) are allowed.	spark_job
Program Path	Path of the program package to be executed. You can enter the path or click **HDFS** or **OBS** to select a file. • The path can contain a maximum of 1,023 characters. It cannot contain special characters (;|&>,<'$) and cannot be left blank or all spaces. • The OBS program path should start with **obs://**, for example, **obs://wordcount/program/*XXX*.jar**. The HDFS program path should start with **hdfs://**, for example, **hdfs://hacluster/user/*XXX*.jar**. • The SparkSubmit job execution program must end with **.jar** or **.py**.	obs://mrs-demotest/program/spark-examples_*XXX*.jar
Tunning Parameters	(Optional) Used to configure optimization parameters such as threads, memory, and vCPUs for the job to optimize resource usage and improve job execution performance. **Table 7-3** lists the common program parameters of Spark jobs. You can configure the parameters based on the execution program and cluster resources.	--class org.apache.spark.examples.SparkPi
Runtime Parameters	(Optional) Key parameters for program execution. The parameters are specified by the function of the user's program. MRS is only responsible for loading the parameters. Multiple parameters are separated by spaces. The value can contain a maximum of 150,000 characters and can be left blank. The value cannot contain special characters such as ;|&><'$ **CAUTION** When entering a parameter containing sensitive information (for example, login password), you can add an at sign (@) before the parameter name to encrypt the parameter value. This prevents the sensitive information from being persisted in plaintext. When you view job information on the MRS console, the sensitive information is displayed as *****. Example: **username=***testuser* **@password=***User password*	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
Service Parameter	(Optional) Service parameters for the job. To modify the current job, change this parameter. For permanent changes to the entire cluster, refer to **Modifying the Configuration Parameters of an MRS Cluster Component** and modify the cluster component parameters accordingly. For example, if decoupled storage and compute is not configured for the MRS cluster and jobs need to access OBS using AK/SK, you can add the following service parameters: • **fs.obs.access.key**: key ID for accessing OBS. • **fs.obs.secret.key**: key corresponding to the key ID for accessing OBS.	-
Command Reference	Commands submitted to the background for execution when a job is submitted.	N/A

Table 7-3 Spark job running program parameters

Parameter	Description	Example
--conf	Spark job configuration in *property=value* format. For more information about parameters required for submitting Spark jobs, see **https://spark.apache.org/docs/latest/submitting-applications.html**.	spark.executor.memory=2G
--conf spark-yarn.maxAppAttempts	Maximum number of retry attempts YARN will make to run the Spark application before giving up. For example, if this parameter is set to **0**, retry is not allowed. If this parameter is set to **1**, one retry is allowed.	0
--driver-memory	Amount of memory to allocate to the driver process in a Spark job. For example, if the allocated memory is 2 GB, set this parameter to **2g**. If the allocated memory is 512 MB, set this parameter to **512m**.	2g
--num-executors	Number of executors to run in a Spark job.	5

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
--executor-cores	Number of CPU cores to use on each executor in a Spark job.	2
--executor-memory	Amount of memory to use per executor in a Spark job. For example, if the allocated memory is 2 GB, set this parameter to **2g**. If the allocated memory is 512 MB, set this parameter to **512m**.	2g
--class	Name of the main class to use for a job. It is specified by a function in your program and is usually a class containing the main method.	org.apache.spark.examples.SparkPi
--files	Upload files to a job. The files can be user-defined configuration files or some data files from OBS or HDFS.	-
--jars	JAR files to include on the driver and executor classpaths. Use commas (,) to separate multiple JAR files.	-

Step 7 Confirm job configuration information and click **OK**.

Step 8 After the job is submitted, you can view the job running status and execution result in the job list. After the job status changes to **Completed**, you can view the analysis result of related programs.

In this sample program, you can click **View Log** to view the detailed execution process of the Spark job.

During job execution, you can click **View Log** or choose **More** > **View Details** to view program execution details. If the job execution is abnormal or fails, you can locate the fault based on the error information.

A created job cannot be modified. If you need to execute the job again, you can click **Clone** to quickly copy the created job and adjust required parameters.

----End

Submitting a Job Using the Cluster Client

Step 1 Prepare the application and data.

This section uses the spark-examples sample program provided by the cluster as an example. The sample program provides multiple functions. You can obtain the sample program from the MRS cluster client (*Client installation directory*/**Spark/spark/examples/jars/spark-examples-***XXX***.jar**). In some versions, the name of the **Spark2x** folder in the cluster is **Spark**. Replace it with the actual name. Upload the obtained sample program to a specified directory in HDFS or OBS. For details, see **Uploading Application Data to an MRS Cluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Program class name: It is specified by a function in your program. In this application, the class name is **SparkPi** for calculating the value of Pi (π).

Step 2 If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.

Step 3 Install an MRS cluster client.

For details, see **Installing an MRS Cluster Client**.

The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

Step 4 Log in to the node where the client is located as the MRS cluster client installation user.

For details, see **Logging In to an MRS Cluster Node**.

Step 5 Run the following command to go to the client installation directory:

```
cd /opt/Bigdata/client
```

Run the following command to load the environment variables:

```
source bigdata_env
```

If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, you do not need to run the **kinit** command.

```
kinit testuser
```

Step 6 Run the following command to switch to the Spark component directory:

```
cd $SPARK_HOME
```

Step 7 Run the following command to submit a Spark job:

```
./bin/spark-submit --master yarn --deploy-mode client --class org.apache.spark.examples.SparkPi examples/jars/spark-examples_*.jar 10
```

The command output is as follows:

```
...
Pi is roughly 3.1402231402231404
```

spark-submit parameter description:

- **deploy-mode**: running mode of the Spark driver. The value can be **client** or **cluster**. In this example, the client mode is used to submit the job.
- **class**: main class name of the application, which is specified by the running application.
- *XXX*.**jar**: program run by the Spark job
- **--conf**: additional configuration properties of the Spark job. If the keytab file is used for user authentication, you can set the following parameters:
  - **spark.yarn.principal**: name of the user who submits the job.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **spark.yarn.keytab**: **keytab** file for user authentication.

For more information about parameters required for submitting Spark jobs, see **https://spark.apache.org/docs/latest/submitting-applications.html**.

Step 8 Log in to FusionInsight Manager as user **testuser**, choose **Cluster** > **Services** > **Yarn**, and click the hyperlink on the right of **ResourceManager Web UI** to access the YARN Web UI. Click the application ID of the job to view the job running information and related logs.

----End

7.3.3 Running a Hive SQL Job

A Hive SQL job is a data analytics task that is written and run based on the Hive data warehouse tool. It is mainly used to process large-scale structured or semi-structured data in a Hadoop distributed cluster. Hive SQL is a SQL dialect defined by Hive. Its syntax is similar to standard SQL, but it is optimized for big data scenarios. Hive SQL jobs can convert SQL statements into MapReduce or Spark jobs for execution, making Hive SQL suitable for non-programmers to quickly perform data query, statistics, and analysis.

MRS allows you to submit and run your own programs, and get the results. This section will show you how to submit a Hive SQL job in an MRS cluster.

A Hive SQL job supports both SQL statements and scripts. If SQL statements contain sensitive information, you can submit the job using a script.

You can create a job online and submit it for running on the MRS console, or submit a job in CLI mode on the MRS cluster client.

Prerequisites

- You have uploaded the program packages and data files required by jobs to OBS or HDFS.
- If the job program needs to read and analyze data in the OBS file system, you need to configure storage-compute decoupling for the MRS cluster. For details, see **Configuring Storage-Compute Decoupling for an MRS Cluster**.

Notes and Constraints

- When the policy of the user group to which an IAM user belongs changes from **MRS ReadOnlyAccess** to **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator**, or vice versa, wait for five minutes after user synchronization for the System Security Services Daemon (SSSD) cache of the cluster node to refresh. Submit a job on the MRS console after the new policy takes effect. Otherwise, the job submission may fail.
- If the IAM username contains spaces (for example, **admin 01**), you cannot create jobs on the MRS console.

Submitting a Job

You can create and run jobs online using the management console or submit jobs by running commands on the cluster client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Submitting a Job on the Console

Step 1 Develop an SQL script.

A Hive SQL job can execute SQL statements directly or execute an SQL script file.

In this section, the script for creating a database and data table, and inserting data in Hive is used as an example. Upload the developed **hive_basic.sql** script file to a specified directory on HDFS, OBS, or an MRS cluster node. For details, see **Uploading Application Data to an MRS Cluster**.

```
-- Create a database.
CREATE DATABASE IF NOT EXISTS test_db;
-- Use the database.
USE test_db;
-- Create a data table.
CREATE TABLE IF NOT EXISTS test_table (
    id INT,
    name STRING,
    age INT,
    gender STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';
-- Insert a single row of data.
INSERT INTO test_table VALUES (1, 'Zhangsan', 30, 'Male');
```

Step 2 Log in to the MRS console.

Step 3 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 4 On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.

Perform this step only when Kerberos authentication is enabled for the cluster.

After IAM user synchronization, wait for five minutes before submitting a job. For details about IAM user synchronization, see Synchronizing IAM Users to MRS.

Step 5 Click **Job Management**. On the displayed job list page, click **Create**.

Step 6 Set **Type** to **HiveSql** and configure Hive SQL job information by referring to **Table 7-4**.

Table 7-4 Job parameters

Parameter	Description	Example
Name	Job name. It can contain 1 to 64 characters. Only letters, digits, hyphens (-), and underscores (_) are allowed.	hive_job
SQL Type	Submission type of the SQL statement. • SQL: Run the entered SQL statement. • Script: Load SQL scripts from HDFS or OBS to run SQL statements.	Script

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
SQL Statement	This parameter is valid only when **SQL Type** is set to **SQL**. Enter the SQL statement to be executed, and then click **Check** to check whether the SQL statement is correct. If you want to submit and execute multiple statements at the same time, use semicolons (;) to separate them.	-
SQL File	This parameter is available only when **SQL Type** is set to **Script**. It specifies the path of the SQL script file to be executed. You can enter the path or click **HDFS** or **OBS** to select a file. • The path can contain a maximum of 1,023 characters. It cannot contain special characters (;|&>,<'$) and cannot be left blank or all spaces. • The OBS program path starts with **obs://**. The HDFS program path starts with **hdfs://hacluster**, for example, **hdfs://hacluster/user/*XXX*.jar**. • The HiveScript file must end with **.sql**.	obs://mrs-demotest/program/hive_basic.sql
Tunning Parameters	(Optional) Used to configure optimization parameters such as threads, memory, and vCPUs for the job to optimize resource usage and improve job execution performance. **Table 7-5** lists the common program parameters of Hive SQL jobs. You can configure the parameters based on the execution program and cluster resources. If you do not configure the parameters, the default values of the cluster are used.	-
Service Parameter	(Optional) Service parameters for the job. To modify the current job, change this parameter. For permanent changes to the entire cluster, refer to **Modifying the Configuration Parameters of an MRS Cluster Component** and modify the cluster component parameters accordingly. For example, if decoupled storage and compute is not configured for the MRS cluster and jobs need to access OBS using AK/SK, you can add the following service parameters: • **fs.obs.access.key**: key ID for accessing OBS. • **fs.obs.secret.key**: key corresponding to the key ID for accessing OBS.	-
Command Reference	Commands submitted to the background for execution when a job is submitted.	N/A

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 7-5 HiveSQL job running program parameters

Parameter	Description	Example Value
--hiveconf	Configure the Hive service.	For example, you can set **hive.execution.engine** to **mr** or **tez** to specify the execution engine. • Parameter: **--hiveconf** • Value: **hive.execution.engine=mr**
--hivevar	Set user-defined variables.	Set the variable ID as follows: • Parameter: **--hivevar id** • Value: **"123" select * from test where id = ${hivevar:id};**

Step 7 Confirm job configuration information and click **OK**.

Step 8 After the job is submitted, you can view the job running status and execution result in the job list. After the job status changes to **Completed**, you can view the analysis result of related programs.

In this sample program, you can click **View Log** to view the detailed execution process of the Hive job.

After the job is successfully executed, connect to Hive Beeline through the client to view the Hive table data.

1. If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

   In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.

2. Install an MRS cluster client.

   For details, see **Installing an MRS Cluster Client**.

   The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Run the following command to go to the client installation directory:

   ```
   cd /opt/Bigdata/client
   ```

   Run the following command to load the environment variables:

   ```
   source bigdata_env
   ```

   If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, you do not need to run the **kinit** command.

   ```
   kinit testuser
   ```

4. Run the following command to access Hive Beeline and view the generated data:

   ```
   beeline
   ```

   Run the following command to view the generated Hive table data:

   ```
   select * from test_db.test_table;
   ```

   Figure 7-3 Viewing Hive table data

   ```
   +----------------+------------------+-----------------+--------------------+
   | test_table.id  | test_table.name  | test_table.age  | test_table.gender  |
   +----------------+------------------+-----------------+--------------------+
   | 1              | Zhangsan         | 30              | Male               |
   +----------------+------------------+-----------------+--------------------+
   ```

----End

Submitting a Job Using the Cluster Client

Step 1 Develop an SQL script.

A Hive SQL job can execute SQL statements directly or execute an SQL script file.

In this section, the script for creating a database and data table, and inserting data in Hive is used as an example. Upload the developed **hive_basic.sql** script file to a specified directory on HDFS, OBS, or an MRS cluster node. For details, see **Uploading Application Data to an MRS Cluster**.

```
-- Create a database.
CREATE DATABASE IF NOT EXISTS test_db;
-- Use the database.
USE test_db;
-- Create a data table.
CREATE TABLE IF NOT EXISTS test_table (
    id INT,
    name STRING,
    age INT,
    gender STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';
-- Insert a single row of data.
INSERT INTO test_table VALUES (1, 'Zhangsan', 30, 'Male');
```

Step 2 If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.

Step 3 Install an MRS cluster client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details, see **Installing an MRS Cluster Client**.

The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

Step 4 Log in to the node where the client is located as the MRS cluster client installation user.

For details, see **Logging In to an MRS Cluster Node**.

Step 5 Run the following command to go to the client installation directory:

```
cd /opt/Bigdata/client
```

Run the following command to load the environment variables:

```
source bigdata_env
```

If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, you do not need to run the **kinit** command.

```
kinit testuser
```

Step 6 Run a **beeline** command to connect to Hive and execute the related job.

```
beeline -f /opt/hive_basic.sql
```

After running the **beeline** command, you can directly execute SQL statements in the Beeline.

```
beeline
```

After the script is executed, view the generated Hive table data in Beeline.

```
select * from test_db.test_table;
```

Figure 7-4 Viewing Hive table data

```
+-----------------+-------------------+------------------+---------------------+
| test_table.id   | test_table.name   | test_table.age   | test_table.gender   |
+-----------------+-------------------+------------------+---------------------+
| 1               | Zhangsan          | 30               | Male                |
+-----------------+-------------------+------------------+---------------------+
```

Use the command below to specify a component service user for clusters with Kerberos authentication disabled. If no service user is specified, HiveServer will be connected using the current OS user.

```
beeline -nMRS cluster service user
```

Step 7 For jobs submitted to YARN for execution, log in to FusionInsight Manager as user **testuser**, choose **Cluster** > **Services** > **Yarn**, click the hyperlink on the right of **ResourceManager Web UI** to access the YARN Web UI, and click the application ID of a job to view the job running information and related logs.

----End

7.3.4 Running a Spark SQL Job

Spark SQL is a Spark module for structured data processing. It integrates relational SQL queries with the distributed computing capabilities of Spark and allows you

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

to perform operations on data using SQL statements or DataFrame/Dataset APIs. Spark SQL jobs are characterized by high performance, ease of use, and high scalability, making them suitable for batch processing, stream processing (such as Structured Streaming), and interactive queries.

MRS allows you to submit and run your own programs, and get the results. This section will show you how to submit a Spark SQL job in an MRS cluster.

A Spark SQL job supports both SQL statements and scripts. If SQL statements contain sensitive information, you can submit the job using a script.

You can create a job online and submit it for running on the MRS console, or submit a job in CLI mode on the MRS cluster client.

Prerequisites

- You have uploaded the program packages and data files required by jobs to OBS or HDFS.
- If the job program needs to read and analyze data in the OBS file system, you need to configure storage-compute decoupling for the MRS cluster. For details, see **Configuring Storage-Compute Decoupling for an MRS Cluster**.

Notes and Constraints

- When the policy of the user group to which an IAM user belongs changes from **MRS ReadOnlyAccess** to **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator**, or vice versa, wait for five minutes after user synchronization for the System Security Services Daemon (SSSD) cache of the cluster node to refresh. Submit a job on the MRS console after the new policy takes effect. Otherwise, the job submission may fail.
- If the IAM username contains spaces (for example, **admin 01**), you cannot create jobs on the MRS console.

Submitting a Job

You can create and run jobs online using the management console or submit jobs by running commands on the cluster client.

Submitting a Job on the Console

Step 1 Develop an SQL script.

A Spark SQL job can execute SQL statements directly or execute an SQL script file.

In this section, the script for creating a Hive data table and inserting data in Spark is used as an example. Upload the developed **spark_basic.sql** script file to a specified directory on HDFS, OBS, or an MRS cluster node. For details, see **Uploading Application Data to an MRS Cluster**.

```
-- Create a table.
CREATE TABLE IF NOT EXISTS test_table1 (
   id INT,
   name STRING,
   age INT,
   gender STRING
)
USING parquet
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
OPTIONS (
    path '/user/hive/warehouse/test_table1',
    compression 'snappy'
);
-- Insert a single row of data.
INSERT INTO test_table1 VALUES (1, 'lisi', 30, 'Male');
```

Step 2 Log in to the MRS console.

Step 3 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 4 On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.

Perform this step only when Kerberos authentication is enabled for the cluster.

After IAM user synchronization, wait for five minutes before submitting a job. For details about IAM user synchronization, see Synchronizing IAM Users to MRS.

Step 5 Click **Job Management**. On the displayed job list page, click **Create**.

Step 6 Set **Type** to **SparkSql** and configure Spark SQL information by referring to **Table 7-6**.

Table 7-6 Job parameters

Parameter	Description	Example
Name	Job name. It can contain 1 to 64 characters. Only letters, digits, hyphens (-), and underscores (_) are allowed.	spark_sql_job
SQL Type	Submission type of the SQL statement • SQL: Run the entered SQL statement. • Script: Load SQL scripts from HDFS or OBS to run SQL statements.	Script
SQL Statement	This parameter is valid only when **SQL Type** is set to **SQL**. Enter the SQL statement to be executed, and then click **Check** to check whether the SQL statement is correct. If you want to submit and execute multiple statements at the same time, use semicolons (;) to separate them.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
SQL File	This parameter is available only when **SQL Type** is set to **Script**. It specifies the path of the SQL script file to be executed. You can enter the path or click **HDFS** or **OBS** to select a file. • The path can contain a maximum of 1,023 characters. It cannot contain special characters (;|&>,<'$) and cannot be left blank or all spaces. • The OBS program path starts with **obs://**. The HDFS program path starts with **hdfs://hacluster**, for example, **hdfs://hacluster/user/*XXX*.jar**. • The script file must end with **.sql**.	obs://mrs-demotest/program/spark_basic.sql
Tunning Parameters	(Optional) Used to configure optimization parameters such as threads, memory, and vCPUs for the job to optimize resource usage and improve job execution performance. **Table 7-7** lists the common program parameters of Spark SQL jobs. You can configure the parameters based on the execution program and cluster resources. If you do not configure the parameters, the default values of the cluster are used.	-
Service Parameter	(Optional) Service parameters for the job. To modify the current job, change this parameter. For permanent changes to the entire cluster, refer to **Modifying the Configuration Parameters of an MRS Cluster Component** and modify the cluster component parameters accordingly. For example, if decoupled storage and compute is not configured for the MRS cluster and jobs need to access OBS using AK/SK, you can add the following service parameters: • **fs.obs.access.key**: key ID for accessing OBS. • **fs.obs.secret.key**: key corresponding to the key ID for accessing OBS.	-
Command Reference	Commands submitted to the background for execution when a job is submitted.	N/A

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 7-7 Program parameters

Parameter	Description	Example
--conf	Spark job configuration in *property=value* format. For more information about parameters required for submitting Spark jobs, see **https://spark.apache.org/docs/latest/submitting-applications.html**.	spark.executor.memory=2G
--driver-memory	Amount of memory to allocate to the driver process in a Spark job. For example, if the allocated memory is 2 GB, set this parameter to **2g**. If the allocated memory is 512 MB, set this parameter to **512m**.	2g
--num-executors	Number of executors to run in a Spark job.	5
--executor-cores	Number of CPU cores to use on each executor in a Spark job.	2
--executor-memory	Amount of memory to use per executor in a Spark job. For example, if the allocated memory is 2 GB, set this parameter to **2g**. If the allocated memory is 512 MB, set this parameter to **512m**.	2g
--jars	Additional dependency packages of a task, which is used to add the external dependency packages to the task.	-

Step 7 Confirm job configuration information and click **OK**.

Step 8 After the job is submitted, you can view the job running status and execution result in the job list. After the job status changes to **Completed**, you can view the analysis result of related programs.

In this sample program, you can click **View Log** to view the detailed execution process of the Spark SQL job.

After the job is successfully executed, connect to Hive Beeline through the client to view the Hive table data.

1. If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

   In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Install an MRS cluster client.

   For details, see **Installing an MRS Cluster Client**.

   The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

3. Run the following command to go to the client installation directory:

   ```
   cd /opt/Bigdata/client
   ```

   Run the following command to load the environment variables:

   ```
   source bigdata_env
   ```

   If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, you do not need to run the **kinit** command.

   ```
   kinit testuser
   ```

4. Run the following command to access Hive Beeline and view the generated data:

   ```
   beeline
   ```

   Run the following command to view the generated Hive table data:

   ```
   select * from test_table1;
   ```

   Figure 7-5 Viewing Hive table data

   ```
   +-----------------+-------------------+------------------+---------------------+
   | test_table1.id  | test_table1.name  | test_table1.age  | test_table1.gender  |
   +-----------------+-------------------+------------------+---------------------+
   | 1               | lisi              | 30               | Male                |
   +-----------------+-------------------+------------------+---------------------+
   ```

----End

Submitting a Job Using the Cluster Client

Step 1 Develop an SQL script.

A Spark SQL job can execute SQL statements directly or execute an SQL script file.

In this section, the script for creating a Hive data table and inserting data in Spark is used as an example. Upload the developed **spark_basic.sql** script file to a specified directory on HDFS, OBS, or an MRS cluster node. For details, see **Uploading Application Data to an MRS Cluster**.

```
-- Create a table.
CREATE TABLE IF NOT EXISTS test_table1 (
    id INT,
    name STRING,
    age INT,
    gender STRING
)
USING parquet
OPTIONS (
    path '/user/hive/warehouse/test_table1',
    compression 'snappy'
);
-- Insert a single row of data.
INSERT INTO test_table1 VALUES (1, 'lisi', 30, 'Male');
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.

Step 3 Install an MRS cluster client.

For details, see **Installing an MRS Cluster Client**.

The MRS cluster comes with a client installed for job submission by default, which can also be used directly. In MRS 3.*x* or later, the default client installation path is **/opt/Bigdata/client** on the Master node. In versions earlier than MRS 3.*x*, the default client installation path is **/opt/client** on the Master node.

Step 4 Log in to the node where the client is located as the MRS cluster client installation user.

For details, see **Logging In to an MRS Cluster Node**.

Step 5 Run the following command to go to the client installation directory:

```
cd /opt/Bigdata/client
```

Run the following command to load the environment variables:

```
source bigdata_env
```

If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, you do not need to run the **kinit** command.

```
kinit testuser
```

Step 6 Run the following command to switch to the Spark component directory:

```
cd $SPARK_HOME
```

Step 7 Run a **spark-sql** command to execute the SQL script.

```
./bin/spark-sql -f /opt/spark_basic.sql
```

After running the **spark-sql** command, you can directly execute SQL statements.

```
spark-sql
```

After the script is executed, view the generated Hive table data.

```
select * from test_table1;
```

Figure 7-6 Viewing table data

```
spark-sql> select * from test_table1;
1       lisi    30      Male
```

Step 8 Log in to FusionInsight Manager as user **testuser**, choose **Cluster** > **Services** > **Spark** or **Spark2x**, click the hyperlink on the right of **Spark WebUI** to access the Spark Web UI, and view Spark job running information and event logs.

----End

7.3.5 Running a Flink Job

Flink jobs are distributed data processing tasks developed based on the Flink framework. They are mainly used for stream data processing and stateful

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

computations. Flink is a unified framework for computations over unbounded (streams) and bounded (batches) data. Flink is designed with stream processing as its core, and it handles batch processing (finite data streams) as a special case of stream processing. Flink features real-time data processing with high throughput, low latency, and exactly-once semantics. It is widely used in real-time monitoring, log analysis, and financial transactions.

MRS allows you to submit and run your own programs, and get the results. This section will show you how to submit a Flink job in an MRS cluster.

You can create a job online and submit it for running on the MRS console, or submit a job in CLI mode on the MRS cluster client.

Prerequisites

- You have uploaded the program packages and data files required by jobs to OBS or HDFS.
- If the job program needs to read and analyze data in the OBS file system, you need to configure storage-compute decoupling for the MRS cluster. For details, see **Configuring Storage-Compute Decoupling for an MRS Cluster**.

Notes and Constraints

- When the policy of the user group to which an IAM user belongs changes from **MRS ReadOnlyAccess** to **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator**, or vice versa, wait for five minutes after user synchronization for the System Security Services Daemon (SSSD) cache of the cluster node to refresh. Submit a job on the MRS console after the new policy takes effect. Otherwise, the job submission may fail.
- If the IAM username contains spaces (for example, **admin 01**), you cannot create jobs on the MRS console.

Submitting a Job

You can create and run jobs online using the management console or submit jobs by running commands on the cluster client.

Submitting a Job on the Console

Step 1 Prepare the application and data.

This section uses the Flink word count application as an example. You can obtain the sample program from the MRS cluster client (*Client installation directory*/**Flink/flink/examples/batch/WordCount.jar**) and upload the program to a specified directory in HDFS or OBS. For details, see **Uploading Application Data to an MRS Cluster**.

To run the application, you need to specify the following parameters:

- **input**: data file to be analyzed, which must be uploaded to the HDFS or OBS file system in advance.

  For example, upload data file **data2.txt**. The file content is as follows:

  ```
  This is a test demo for MRS Flink. Flink is a unified computing framework that supports both batch processing and stream processing. It provides a stream data processing engine that supports data distribution and parallel computing.
  ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **output**: file for storing the word count result.

Step 2 Log in to the MRS console.

Step 3 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 4 On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.

Perform this step only when Kerberos authentication is enabled for the cluster.

After IAM user synchronization, wait for five minutes before submitting a job. For details about IAM user synchronization, see Synchronizing IAM Users to MRS.

Step 5 Click **Job Management**. On the displayed job list page, click **Create**.

Step 6 Set **Type** to **Flink** and configure Flink job information by referring to **Table 7-8**.

Table 7-8 Job parameters

Parameter	Description	Example
Name	Job name. It can contain 1 to 64 characters. Only letters, digits, hyphens (-), and underscores (_) are allowed.	flink_job
Program Path	Path of the program package to be executed. You can enter the path or click **HDFS** or **OBS** to select a file. • The path can contain a maximum of 1,023 characters. It cannot contain special characters (;|&>,<'$) and cannot be left blank or all spaces. • The OBS program path starts with **obs://**. The HDFS program path starts with **hdfs://hacluster**, for example, **hdfs://hacluster/user/*XXX*.jar**. • The Flink job execution program must end with **.jar**.	obs://mrs-demotest/program/WordCount.jar
Tunning Parameters	(Optional) Used to configure optimization parameters such as threads, memory, and vCPUs for the job to optimize resource usage and improve job execution performance. **Table 7-9** lists the common program parameters of Flink jobs. You can configure the parameters based on the execution program and cluster resources. If you do not configure the parameters, the default values of the cluster are used.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
Runtime Parameters	(Optional) Key parameters for program execution. The parameters are specified by the function of the user's program. MRS is only responsible for loading the parameters. Multiple parameters are separated by spaces. The value can contain a maximum of 150,000 characters and can be left blank. The value cannot contain special characters such as ;|&><'$ **CAUTION** When entering a parameter containing sensitive information (for example, login password), you can add an at sign (@) before the parameter name to encrypt the parameter value. This prevents the sensitive information from being persisted in plaintext. When you view job information on the MRS console, the sensitive information is displayed as *****. Example: **username=***testuser* **@password=***User password*	--input obs://mrs-demotest/input/data2.txt --output obs://mrs-demotest/output.txt
Service Parameter	(Optional) Service parameters for the job. To modify the current job, change this parameter. For permanent changes to the entire cluster, refer to **Modifying the Configuration Parameters of an MRS Cluster Component** and modify the cluster component parameters accordingly. For example, if decoupled storage and compute is not configured for the MRS cluster and jobs need to access OBS using AK/SK, you can add the following service parameters: • **fs.obs.access.key**: key ID for accessing OBS. • **fs.obs.secret.key**: key corresponding to the key ID for accessing OBS.	-
Command Reference	Commands submitted to the background for execution when a job is submitted.	N/A

Table 7-9 Parameters for running a Flink job

Parameter	Description	Example
-ytm	Memory size of each TaskManager container. (You can select a unit as required, and the default value is MB.)	1024
-yjm	Memory size of JobManager container. (You can select a unit as required, and the default value is MB.)	1024

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example
-ys	Number of TaskManager cores.	2
-ynm	Custom name of an application on YARN.	test
-c	Class of the program entry method (for example, the **main** or **getPlan()** method). This parameter is required only when the JAR program does not specify the class of its manifest.	com.bigdata.mrs.test

Step 7 Confirm job configuration information and click **OK**.

Step 8 After the job is submitted, you can view the job running status and execution result in the job list. After the job status changes to **Completed**, you can view the analysis result of related programs.

In this sample program, you can click **View Log** to view the detailed execution process of the Flink job.

After the job is executed, you can view the word count result in the specified result output directory.

For example, the result is stored in the **obs://mrs-demotest/output.txt** file, and the file content is as follows:

```
a 3
and 2
batch 1
both 1
computing 2
data 2
demo 1
distribution 1
engine 1
flink 2
for 1
framework 1
is 2
it 1
mrs 1
parallel 1
processing 3
provides 1
stream 2
supports 2
test 1
that 2
this 1
unified 1
```

----End

Submitting a Job Using the Cluster Client

Step 1 Prepare the application and data.

This section uses the Flink word count application as an example. You can obtain the sample program from the MRS cluster client (*Client installation directory*/

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Flink/flink/examples/batch/WordCount.jar) and upload the program to a specified directory in HDFS or OBS. For details, see **Uploading Application Data to an MRS Cluster**.

To run the application, you need to specify the following parameters:

- **input**: data file to be analyzed, which must be uploaded to the HDFS or OBS file system in advance.

  For example, upload data file **data2.txt**. The file content is as follows:

  ```
  This is a test demo for MRS Flink. Flink is a unified computing framework that supports both batch processing and stream processing. It provides a stream data processing engine that supports data distribution and parallel computing.
  ```

- **output**: file for storing the word count result.

Step 2 If Kerberos authentication has been enabled for the current cluster, create a service user with job submission permissions on FusionInsight Manager in advance. For details, see **Creating an MRS Cluster User**.

1. In this example, create human-machine user **testuser**, and associate the user with user group **supergroup** and role **System_administrator**.
2. On FusionInsight Manager, choose **System** > **Permission** > **User**. On the displayed page, locate the row that contains the added user, click **More** in the **Operation** column, and select **Download Authentication Credential**.
3. Decompress the downloaded authentication credential package and copy the obtained file to the client node, for example, the **/opt/flinkclient/Flink/flink/conf** directory on the client node. If the client is installed on a node outside the cluster, copy the obtained files to the **/etc/** directory on the node.

Step 3 Install an MRS cluster client.

For details, see **Installing an MRS Cluster Client**. The subsequent operations use the installation directory **/opt/flinkclient** as an example.

Step 4 Log in to the node where the client is located as the MRS cluster client installation user.

For details, see **Logging In to an MRS Cluster Node**.

Step 5 Run the following command to go to the client installation directory:

```
cd /opt/flinkclient
```

Run the following command to load the environment variables:

```
source bigdata_env
```

Step 6 If Kerberos authentication has been enabled for the current cluster, modify Flink security configuration parameters.

1. MRS 3.x or later: In a cluster with Kerberos authentication enabled, you need to **add** the floating IP addresses of the client installation node and cluster's FusionInsight Manager to the **jobmanager.web.allow-access-address** configuration item in the **flink-conf.yaml** file. Use commas (,) to separate the IP addresses.

   ```
   vi Flink/flink/conf/flink-conf.yaml
   ```

   For example, the changes are as follows:

   ```
   jobmanager.web.allow-access-address: 192.168.0.219,192.168.0.153,192.168.0.47,192.168.0.127
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can run the **ifconfig** command on the OMS node of the cluster to view the floating IP address of the cluster's FusionInsight Manager. The value of **eth*X*:wsom** indicates the floating IP address.

2. Search for and configure the keytab path and username in the **flink-conf.yaml** file.

   For example, the changes are as follows:

   ```
   security.kerberos.login.keytab: /opt/flinkclient/Flink/flink/conf/user.keytab
   security.kerberos.login.principal: testuser
   ```

3. Run a command to perform Flink security hardening.

   Run the following command to go to the Flink client directory:

   ```
   cd Flink/flink/bin
   ```

   Execute the following script and set a custom password as prompted. This script automatically replaces the SSL-related configuration values in *Client directory*/**Flink/flink/conf/flink-conf.yaml**. For versions earlier than MRS 3.x, external SSL is disabled by default for clusters in security mode. If you need to enable external SSL, configure the parameters and execute the script again. **Table 7-10** describes the parameters.

   ```
   sh generate_keystore.sh
   ```

   Table 7-10 Parameter description

Parameter	Description	Example Value
security.ssl.rest.enabled	Switch to enable external SSL.	true
security.ssl.rest.keystore	Path for storing **keystore**.	${path}/flink.keystore
security.ssl.rest.keystore-password	Password of the **keystore**. **123456** indicates a user-defined password is required.	123456
security.ssl.rest.key-password	Password of the SSL key. **123456** indicates a user-defined password is required.	123456
security.ssl.rest.truststore	Path for storing the **truststore**.	${path}/flink.truststore
security.ssl.rest.truststore-password	Password of the **truststore**. **123456** indicates a user-defined password is required.	123456

4. Use relative paths for **flink.keystore** and **flink.truststore** files.

   a. Create a new directory, for example, **ssl**, in **Flink/flink/conf/**.

      ```
      mkdir /opt/flinkclient/Flink/flink/conf/ssl
      ```

   b. Move the **flink.keystore** and **flink.truststore** file to the /**opt/Bigdata/client/Flink/flink/conf/ssl/** directory.

      ```
      mv flink.keystore /opt/flinkclient/Flink/flink/conf/ssl/
      mv flink.truststore /opt/flinkclient/Flink/flink/conf/ssl/
      ```

   c. For MRS 3.*x* or later: Change the values of the following parameters in the **flink-conf.yaml** file to relative paths:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
security.ssl.keystore: ssl/flink.keystore
security.ssl.truststore: ssl/flink.truststore
```

d. For MRS 3.x or earlier: Change the values of the following parameters in the **flink-conf.yaml** file to relative paths:

```
security.ssl.internal.keystore: ssl/flink.keystore
security.ssl.internal.truststore: ssl/flink.truststore
```

5. If the cluster client is installed on a node outside the cluster, add the following configurations to the **flink-conf.yaml** file to configure the IP address of the node where the client is located:

```
web.access-control-allow-origin: 192.168.0.150
jobmanager.web.allow-access-address: 192.168.0.150
```

Step 7 Run the Flink job.

This section uses the word count sample program provided by the client as an example.

- Security-mode cluster (Kerberos authentication enabled)
  - Run the command in the directory at the same level as **ssl** to start the session and submit the job in the session.

    For example, if the **ssl** directory is **/opt/flinkclient/Flink/flink/conf/**, run the following command in the **/opt/flinkclient/Flink/flink/conf/** directory:

    ```
    yarn-session.sh -t ssl/ -nm "session-name"
    ```

    Open another client connection window and run the following command to initialize environment variables:

    ```
    source /opt/flinkclient/bigdata_env
    ```

    Run the following command to submit the job:

    ```
    flink run /opt/flinkclient/Flink/flink/examples/batch/WordCount.jar --input obs://mrs-demotest/input/data2.txt --output obs://mrs-demotest/output.txt
    ```

  - Run the following command to submit a single job on YARN:

    ```
    flink run -m yarn-cluster /opt/flinkclient/Flink/flink/examples/batch/WordCount.jar --input obs://mrs-demotest/input/data2.txt --output obs://mrs-demotest/output.txt
    ```

- Normal-mode cluster (Kerberos authentication disabled)
  - Run the following commands in the **/opt/flinkclient/Flink/flink/conf/** directory to start a session and submit the job in the session:

    ```
    yarn-session.sh -nm "session-name" -d
    ```

    Open another client connection window and run the following command to submit the job:

    ```
    flink run /opt/flinkclient/Flink/flink/examples/batch/WordCount.jar --input obs://mrs-demotest/input/data2.txt --output obs://mrs-demotest/output.txt
    ```

  - Run the following command to submit a single job on YARN:

    ```
    flink run -m yarn-cluster /opt/flinkclient/Flink/flink/examples/batch/WordCount.jar --input obs://mrs-demotest/input/data2.txt --output obs://mrs-demotest/output.txt
    ```

Step 8 Log in to FusionInsight Manager as user **testuser**, choose **Cluster** > **Services** > **Yarn**, and click the hyperlink on the right of **ResourceManager Web UI** to access the YARN Web UI. Click the application ID of the job to view the job running information and related logs.

Step 9 After the job is executed, you can view the word count result in the specified result output directory.

For example, the result is stored in the **obs://mrs-demotest/output.txt** file, and the file content is as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
a 3
and 2
batch 1
both 1
computing 2
data 2
demo 1
distribution 1
engine 1
flink 2
for 1
framework 1
is 2
it 1
mrs 1
parallel 1
processing 3
provides 1
stream 2
supports 2
test 1
that 2
this 1
unified 1
```

----End

7.4 Viewing MRS Job Details and Logs

MRS is a service for deploying and managing the Hadoop system on the cloud. You can manage, submit, and execute jobs, and view job status, detailed job configurations, and run logs on the MRS console. This helps you to quickly learn the job running status, optimize job configurations, and rectify faults.

Notes and Constraints

No logs are generated on the backend nodes for Spark SQL and DistCp jobs. Then, you cannot view the run logs of running Spark SQL and DistCp jobs on the console.

Viewing the Job Status

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click **Job Management** to view the list and status of jobs created in an MRS cluster.

By default, jobs in the job list are sorted by submission time. For details about the parameters in the job list, see **Table 7-11**. You can quickly filter jobs by job type and job status.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 7-11 Job list parameters

Parameter	Description
Name/ID	Job name, which is set when a job is created. ID is the unique identifier of a job. After a job is added, the system automatically assigns a value to ID.
Username	Name of the user who submits a job.
Job Type	Job type. After importing and exporting files on the **Files** tab page, you can view the DistCp job on the **Jobs** page.
Status	Job status. • **Submitted**: The job has been submitted. • **Accepted**: Initial status of a job after it is successfully submitted. • **Running**: The job is running. • **Completed**: The job is executed. • **Terminated**: The job execution is stopped. • **Abnormal**: An error is reported during job execution, or the job execution is complete but fails.
Result	Execution result of a job. • **Undefined**: indicates that the job is being executed. • **Successful**: indicates that the job has been executed. • **Killed**: indicates that the job is manually terminated during execution. • **Failed**: indicates that the job fails to be executed. After a job is successfully executed or fails to be executed, it cannot be executed again. You can only add a job again.
Queue Name	Name of the resource queue bound to the user who submits the job.
Submitted	Time when a job is submitted.
Ended	Time when a job is completed or manually stopped.
Operation	Operations can be performed on the job. For details, see **Managing MRS Cluster Jobs**.

----End

Viewing Job Logs on the Console

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters** in the navigation pane, select a running cluster, and click its name to switch to the cluster details page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click **Jobs**.

Step 4 In the **Operation** column of the job to be viewed, choose **More** > **View Details**.

In the **View Details** window that is displayed, configuration of the selected job is displayed.

Step 5 Select a running job, and click **View Log** in the **Operation** column.

In the new page that is displayed, real-time log information of the job is displayed.

- Each tenant can submit and view 10 jobs concurrently.
- After a job is executed, the system will compress and save the logs to the corresponding path if you choose to save job logs to OBS or HDFS. Therefore, after a job execution of this type is completed, the job status is still **Running**. After the log is successfully stored, the job status changes to **Completed**. The log storage duration depends on the log size and takes several minutes.
- When you create, stop, or delete a job in an MRS cluster on the management console or using the APIs, the operation will not be recorded on the **Operation Logs** page. You can check the operation records on the CTS console.

----End

Viewing Job Logs on the YARN Web UI

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > **Yarn** and click the hyperlink on the right of **ResourceManager Web UI**.

Step 3 On the displayed **All Applications** page, click **Logout** in the upper right corner. Then, log in again as a YARN administrator or a user who submits jobs.

Step 4 On the **All Applications** page, click the application ID of a job to view the job running information, and click **Logs** to view the logs.

Step 5 View the logs of a specified YARN job on the client.

1. Log in to the node where the cluster client is located as user **root**.
2. Run the following command to initialize environment variables:
   source *Client installation directory*/bigdata_env
3. If Kerberos authentication is enabled for the current cluster, run the following command to authenticate the current user (YARN administrator or a user who submits jobs). If Kerberos authentication is disabled for the current cluster, skip this step.
   kinit *MRS cluster user*
4. Run the following command to obtain the logs of a specified job:
   yarn logs -applicationId *Application ID of the required job*

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Helpful Links

You can manually stop a running MRS job or delete a completed job on the management console. For details, see **Stopping and Deleting an MRS Cluster Job**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8 Managing Clusters

8.1 MRS Cluster Overview

MRS Console

After creating a cluster, you can use the MRS console or MRS Manager to manage and maintain it, as well as view basic information.

- The MRS console's cluster management page allows you to monitor and manage nodes, components, alarms, files, and jobs related to the cluster.
- MRS Manager is the O&M management system of MRS and provides unified cluster management capabilities for services deployed in clusters.

For further information on the capabilities of the MRS console and Manager, refer to **Table 8-1**.

Table 8-1 MRS console and Manager capability comparison

Operation	Console	Manager
Changing subnets, adding security group rules, controlling OBS permissions, managing agencies, and synchronizing IAM users	Yes	No
Adding node groups, scaling out, scaling in, and upgrading specifications	Yes	No
Isolating hosts, starting all roles, and stopping all roles	Yes	Yes
Downloading the client, starting services, stopping services, and perform rolling restart of services	Yes	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation	Console	Manager
Viewing the instance status of services, configuring parameters, and synchronizing configurations	Yes	Yes
Viewing cleared alarms and events	Yes	Yes
Viewing the alarm help	Yes	Yes
Setting thresholds for threshold alarms	No	Yes
Adding message subscription specifications	Yes	No
Managing files	Yes	No
Managing jobs	Yes	No
Managing tenants	Yes	Yes
Managing tags	Yes	No
Setting permissions (adding and deleting users, user groups, and roles)	No	Yes
Backing up component data and restoring data	No	Yes
Logging metadata operation audit	No	Yes
Monitoring resources	Yes	Yes

MRS Cluster Management Objects

There are various fundamental objects in MRS, which are listed in **Table 8-2**.

Table 8-2 MRS basic objects

Object	Description	Example
Component	Function set that can complete specific business.	KrbServer and LdapServer
Instance	Specific instance of a component, which can also be referred to as a component.	KrbServer
Role	Function entity that forms a complete component, usually called role.	KrbServer is composed of the KerberosAdmin role and KerberosServer role.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Object	Description	Example
Role instance	Specific instance of a service role running on a host.	KerberosAdmin that is running on Host2 and KerberosServer that is running on Host3
Host	An ECS running Linux OS.	Host1 to Host5
Rack	Physical entity that contains multiple hosts connecting to the same switch. Physical racks are invisible to ECSs in an MRS cluster. Rack names on MRS Manager represent logical racks.	Rack1 contains Host1 to Host5.
Cluster	Logical entity that consists of multiple hosts and provides various services.	Cluster names **Cluster1** consists of five hosts (Host1 to Host5) and provides services such as KrbServer and LdapServer.

8.2 Accessing MRS Manager

MRS is a one-stop management platform tailored for enterprise-scale big data. By deeply integrating open-source big data components (such as Hadoop, Spark, and Flink), it simplifies deployment and boosts O&M efficiency by providing a standardized O&M system and visualized management and control capabilities. This helps businesses unlock valuable insights from their data faster than ever before.

- Visualized monitoring system

  Displays the cluster health status, resource usage (CPU, memory, and disk), and service running metrics in real time. You can customize monitoring panels and alarm policies to dynamically obtain key system information.

- Operations with a few clicks

  Provides a graphical user interface for cluster, service, and role instance management, and supports operations such as starting, stopping, and configuration adjustment, greatly improving O&M efficiency.

- Enhanced data security

  Integrates fine-grained permission control (such as user- or role-based access control) and operation audit logs to accurately record data access and cluster management operations, meeting enterprise-level data security and compliance requirements.

Currently, you can choose any of the following methods to access MRS Manager of an MRS cluster as required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-3 Methods for accessing MRS Manager

Access Methods	Description	Application Scenario	Constraint
Accessing MRS Manager Using an EIP	Bind an EIP to the MRS cluster so that external users can directly access MRS Manager.	This method is simple and easy to use. You can quickly configure it when creating an MRS cluster. It is recommended.	You need to obtain an EIP and bind it to the cluster.
Accessing MRS Manager Using Direct Connect	Access MRS Manager using a Direct Connect connection.	After a Direct Connect connection has been set up to connect your local network and the MRS cluster network, you can use this method to access MRS Manager.	You need to ensure that Direct Connect is available, and a Direct Connect connection has been established between your local data center and the VPC on the cloud. You can switch between the EIP and Direct Connect access modes on the MRS console.
Accessing MRS Manager Using a Browser on an ECS	Access MRS Manager using a browser on an ECS that can communicate with the MRS cluster network.	Use this method to access MRS Manager when EIPs are not supported.	To use this method, you need to create an ECS in the same VPC and security group as the MRS cluster.
Accessing MRS Manager Through an SSH Tunnel	Create an SSH tunnel connection to send user access requests to the MRS cluster and dynamically forward the requests to the corresponding site.	Use this method to access MRS Manager when the local PC and the MRS cluster are on different networks.	The configurations are complex, and there are high requirements on network bandwidth and connection stability. macOS does not support this method.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- Make sure the cluster is not in the starting, stopping, stopped, deleting, deleted, or frozen state before accessing MRS Manager.
- You have obtained the user information for logging in to MRS Manager from the cluster administrator.

  You can log in to MRS Manager as the default **admin** user or a user created by the administrator with the required login permission.

Accessing MRS Manager of a Cluster

You can access MRS Manager of a cluster using any of the following methods based on your actual requirements and network environment.

Accessing MRS Manager Using an EIP

Step 1 Log in to the MRS console.

Step 2 In the navigation pane, choose **Active Clusters**. Click the target cluster name to access the cluster details page.

Step 3 Click **Access Manager** next to **MRS Manager**. In the displayed dialog box, select **EIP** and configure the EIP information.

1. If no EIP is bound when the MRS cluster is created, select an available EIP from the EIP drop-down list. Otherwise, go to the next step.
   - If no available EIPs are displayed, click **Manage EIP** to create one. An EIP can be bound to only one MRS cluster.
   - To unbind or release an EIP, log in to the **EIPs** page, locate the row containing the target EIP, and click **Unbind** or choose **More** > **Release** in the **Operation** column.
2. In **Security Group**, select the security group to which the current cluster belongs. The security group is configured during cluster creation or is automatically created by the cluster.

   If you want to view, modify, or delete a security group rule, click **Manage Security Group Rule**.
   - "MRS Manager public ip access control rule" is automatically added to the **Description** column of the added inbound rule of the security group for easy identification.
   - It is normal that the automatically generated public IP address is different from your local IP address and no action is required.
   - Port **9022** is the Knox port of the MRS cluster. Therefore, you need to obtain the permission on this port for accessing MRS Manager.
3. Select the information to be confirmed and click **OK**. The MRS Manager login page is displayed.

Step 4 Enter the login information and click **Log In**.

You can log in to MRS Manager as user **admin** or a user created by the administrator with the required login permission.

The password of user **admin** is the custom password set during MRS cluster creation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 (Optional) To grant users in other network segments the permission to access MRS Manager, you can modify the security group and add the IP address range of the users.

1. On the cluster details page, click **Add Security Group Rule** next to **EIP**.
2. In the displayed **Add Security Group Rule** dialog box, add the IP address segment required for authorized users to access the public network, and select the check box.

   By default, the IP address segment used for accessing the public network is filled. You can change the IP address segment as required. To view, modify, or delete security group rules, click **Manage Security Group Rule**.
3. Click **OK**.

----End

Accessing MRS Manager Using Direct Connect

Step 1 Log in to the MRS console.

Step 2 Click the name of a cluster to enter its details page.

Step 3 On the **Dashboard** page of the cluster details page, click **Access Manager** next to **MRS Manager**.

Step 4 Set **Access Mode** to **Direct Connect** and confirm that you understand the impact of the operation.

The floating IP address is automatically allocated by MRS to access MRS Manager. Before using Direct Connect to access MRS Manager, ensure that **the connection between the local data center and the VPC has been established**.

Step 5 Click **OK**. The MRS Manager login page is displayed.

Step 6 Enter the login information and click **Log In**.

You can log in to MRS Manager as user **admin** or a user created by the administrator with the required login permission.

The password of user **admin** is the custom password set during MRS cluster creation.

----End

Accessing MRS Manager Using a Browser on an ECS

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, click the name of the specified cluster.

Record the **AZ**, **VPC**, and **Security Group** of the cluster.

Step 3 On the homepage of the management console, choose **Service List** > **Elastic Cloud Server** to switch to the ECS management console and create an ECS.

- The **AZ**, **VPC**, and **Security Group** of the ECS must be the same as those of the cluster to be accessed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Select a Windows public image.

  For example, a standard image **Windows Server 2012 R2 Standard 64bit(40GB)**.
- For details about other parameters, see the ECS operation guide.

  If the security group of the ECS is different from **Default Security Group** of the Master node, you can modify the configuration using either of the following methods:

  – Change the security group of the ECS to the default Master node security group.

  – Add two security group rules to the security groups of the Master and Core nodes to enable the ECS to access the cluster. Set **Protocol** to **TCP**, **Ports** of the two security group rules to **28443** and **20009**, respectively.

  If "Failed to add security group rules." is displayed, check whether the security group quota is sufficient. If more quotas are needed, increase the quotas or delete security group rules that are no longer used.

Step 4 On the EIP console, apply for an EIP and bind it to the ECS.

Step 5 Log in to the ECS.

The Windows system account, password, EIP, and security group rules are required for logging in to the ECS.

Step 6 On the Windows remote desktop, use your browser to access Manager.

The Manager access address is in the **https://***OMS floating IP address***:28443/web** format.

To query the OMS floating IP address, perform the following steps:

1. Remotely log in to the Master1 node.

   For details about how to log in to a node in the MRS cluster, see **Logging In to an MRS Cluster Node**.
2. Run the following command. In the command output, **eth0:wsom** indicates the OMS floating IP address. Record the value of **inet**.

   ```
   ifconfig
   ```

   If the OMS floating IP address cannot be obtained on the Master1 node, switch to the Master2 node to query and record the floating IP address. If there is only one Master node, query and record the IP address on this node.

Step 7 Enter the login information and click **Log In**.

You can log in to MRS Manager as user **admin** or a user created by the administrator with the required login permission.

The password of user **admin** is the custom password set during MRS cluster creation

----End

Accessing MRS Manager Through an SSH Tunnel

Step 1 Ensure that you have made the following preparations:

- You have prepared an SSH client for creating the SSH tunnel, for example, the Git open source SSH client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You have created an MRS cluster and prepared a key file in PEM format or obtained the password of user **root** used during cluster creation.
- Your local environment can access the Internet.

Step 2 Log in to the MRS console.

Step 3 Choose **Active Clusters** and click a specified MRS cluster.

Record information about the security group of the cluster.

Step 4 Add an inbound rule to the security group of the Master node to allow data access from the MRS cluster IP address through port **22**.

Step 5 Query the primary management node of the cluster by referring to **Checking MRS Active/Standby Management Nodes**.

Step 6 Bind an EIP to the active management node of the cluster.

Step 7 Start Git Bash locally and run the following command to log in to the active management node of the cluster:

```
ssh root@EIP
```

Alternatively, run the following command:

```
ssh -i Key file path root@EIP
```

Step 8 View data forwarding configurations.

```
cat /etc/sysctl.conf | grep net.ipv4.ip_forward
```

- If **net.ipv4.ip_forward=1** is displayed, the forwarding function has been configured. Go to **Step 10**.
- If **net.ipv4.ip_forward=0** is displayed, the forwarding function has not been configured. Go to **Step 9**.
- If **net.ipv4.ip_forward** fails to be queried, this parameter has not been configured. Run the following command and then go to **Step 10**:

```
echo "net.ipv4.ip_forward = 1" >> /etc/sysctl.conf
```

Step 9 Modify forwarding configurations on the node.

1. Switch to user **root**.

```
sudo su - root
```

2. Modify forwarding configurations.

```
echo 1 > /proc/sys/net/ipv4/ip_forward
sed -i "s/net.ipv4.ip_forward=0/net.ipv4.ip_forward = 1/g" /etc/sysctl.conf
sysctl -w net.ipv4.ip_forward=1
```

3. Modify the **sshd** configuration file.

```
vi /etc/ssh/sshd_config
```

Press **I** to enter editing mode. Locate **AllowTcpForwarding** and **GatewayPorts** and delete comment tags. Modify them as follows. Save the changes and exit.

For example, the changes are as follows:

```
AllowTcpForwarding yes
GatewayPorts yes
```

4. Restart the sshd service.

```
service sshd restart
```

Step 10 View the floating IP address.

```
ifconfig
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In the command output, **eth0:FI_HUE** indicates the floating IP address of Hue, and **eth0:wsom** indicates the floating IP address of Manager. Record the value of **inet**.

Run the command to exit.

```
exit
```

Step 11 Run the following command on the local host to create an SSH tunnel that supports dynamic port forwarding:

```
ssh -i Path of the key file -v -ND Local port root@EIP
```

Alternatively, run the following command:

```
ssh -v -ND Local port root@EIP
```

Enter the password of user **root** used during cluster creation as prompted.

In the command, set **Local port** to the user's local port that is not occupied. Port **8157** is recommended.

After the SSH tunnel is created, use **-D** to enable the dynamic port forwarding function.

By default, the dynamic port forwarding function enables a SOCKS proxy process and monitors the user's local port. Port data will be forwarded to the primary management node using the SSH tunnel.

Step 12 Configure the browser proxy.

1. Go to the Google Chrome client installation directory on the local PC.
2. Press **Shift** and right-click the blank area, choose **Open Command Window Here** and enter the following command:

   ```
   chrome --proxy-server="socks5://localhost:8157" --host-resolver-rules="MAP * 0.0.0.0 , EXCLUDE localhost" --user-data-dir=c:/tmppath --proxy-bypass-list="*google*com,*gstatic.com,*gvt*.com,*:80"
   ```

   - In the preceding command, **8157** is the local proxy port configured in **Step 11**.
   - If the local OS is Windows 10, start the Windows OS, click **Start** and enter **cmd**. In the displayed CLI, run the preceding command.

     If this method fails, click **Start**, enter the command in the search box, and run the command.

Step 13 In the address box of the browser, enter the address for accessing Manager.

The Manager access address is in the **https://***Manager floating IP address*:**28443/web** format.

For the first access, add the site to the trusted site list as prompted to continue to open the page.

Step 14 Enter the login information and click **Log In**.

You can log in to MRS Manager as user **admin** or a user created by the administrator with the required login permission.

The password of user **admin** is the custom password set during MRS cluster creation.

Step 15 When logging out of Manager, terminate and close the SSH tunnel.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Helpful Links

- After logging in to MRS Manager, you can view the cluster running status, start or stop the cluster, and synchronize cluster configurations. For details, see **Managing an MRS Cluster**.
- After logging in to MRS Manager, you can view the component running status, start or stop components, and modify component configurations. For details, see **Managing MRS Cluster Components**.
- After logging in to MRS Manager, you can view the cluster monitoring metrics and alarm information. For details, see **Viewing MRS Cluster Monitoring Metrics** and **Viewing and Configuring MRS Alarm Events**.
- After logging in to MRS Manager, you can create and manage MRS cluster users. For details, see **Managing MRS Cluster Users**.
- On MRS Manager, you can access the open-source web UIs of components such as Hadoop and Spark in the MRS cluster. For details, see **Accessing Web UIs of Open-Source Components Managed in MRS Clusters**.

8.3 Introduction to MRS Manager

MRS manages and analyzes massive amounts of data and quickly mines valuable data from structured and unstructured data. Open source components have complex structures and therefore they are difficult to install, configure, and manage. MRS Manager is a unified enterprise-level cluster management platform that provides:

- Cluster monitoring: enables you to quickly learn the running status of hosts and services.
- Graphical metric monitoring and customization: enable you to obtain key system information in a timely manner.
- Service property configurations can meet service performance requirements.
- Cluster, service, and role instance operations: allow you to start or stop services and clusters with just a few clicks.
- Rights management and audit: allow you to configure the access control and manage operation logs.

MRS Manager Homepage

MRS Manager provides a unified cluster management platform, facilitating rapid and easy O&M for clusters.

The MRS Manager GUI contains three parts: the operation bar at the top, the display area in the middle, and the taskbar at the bottom. The operation bar has multiple entries, each with its unique function. These functions are explained in the table below.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-4 Functions of different entries (MRS 3.x)

Entry	Function
Homepage	Displays key monitoring metrics of clusters and host statuses in column charts, line charts, and tables. You can customize a dashboard to display key monitoring metrics and drag and drop it anywhere on the GUI. Data on the dashboard can be automatically refreshed.
Cluster	Provides guidance on how to monitor, operate, and configure services in a cluster, helping you manage services in a unified manner.
Hosts	Provides guidance on how to monitor and operate hosts, helping you manage hosts in a unified manner.
O&M	Provides guidance on how to query and handle alarms, helping you identify and rectify product faults and potential risks in a timely manner to ensure normal system operations.
Audit	Allows you to query and export audit logs, and view all user activities and operations.
Tenant Resources	Provides a unified tenant management platform.
System	Provides system management settings for MRS Manager, such as user permission settings.

On the MRS Manager homepage, you can view the cluster service status, monitoring reports, and alarm statistics and analysis information in the preview area.

- In the upper right corner of the homepage, you can view the number of alarms of different severities, number of running tasks, current user, and help information.

  Click  to view the task name, status, progress, start time, and end time of the last 100 operation tasks in **Task Management Center**.

  For a start, stop, restart, or rolling restart task, you can abort it by clicking the task name in the task list, clicking **Abort**, and then entering the system administrator password in the dialog box that is displayed. An aborted task is no longer executed.
- The taskbar at the bottom of the homepage displays the language options of FusionInsight Manager and the current cluster time and time zone information. You can switch the system language as needed.
- Service status preview area

  The list of installed service components in the cluster is displayed on the left of the homepage. You can view the status and alarms of each service.

  The  icon on the left of each service name indicates that the service is running properly; the  icon indicates that the current service fails to start; and the  icon indicates that the current service is not started.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can also check whether alarms have been generated for the service on the right of the service name. If alarms have been generated, the alarm severities and the number of alarms are displayed.

The icon displayed on the right of the service name indicates that the service configuration has expired.

- Monitoring status report area

  The chart area is on the right of the homepage, which displays key monitoring metric reports, such as the status of all hosts in the cluster, host CPU usage, and host memory usage. You can customize the monitoring reports displayed in the chart area. For details, see **Viewing MRS Cluster Resource Monitoring Metrics**.

  You can view the data source of a monitoring chart in the lower left corner of the chart. You can zoom in on a monitoring report to view chart values more clearly or close the monitoring report.

- Alarm analysis

  On the **Alarm Analysis** page, you can view the **Top 20 Alarms** table and **Analysis on Top 3 Alarms** chart. You can click an alarm name in the **Top 20 Alarms** table to view analysis information of this alarm only. Alarm analysis allows you to view top alarms and their occurrence time so you can handle alarms accordingly, improving system stability.

  NOTE

  For MRS 3.3.0 and later versions, the alarm information page is different from that of other versions.

Checking the MRS Manager Version

By checking the MRS Manager version, you can get ready for system upgrades and regular maintenance.

The following operations use an MRS cluster of 3.*x* or later as an example.

- Using the GUI:

  Log in to MRS Manager of the MRS cluster. On the homepage, click in the upper right corner and choose **About** from the drop-down list. In the dialog box that is displayed, check the MRS Manager version.

- Using the CLI:

  a. Log in to the Manager active OMS node as user **root**.

  To obtain the IP address of the active OMS node, view the host information marked with ★ on the **Hosts** page of Manager.

  b. Check the version and platform information of Manager.

```
su - omm
cd ${BIGDATA_HOME}/om-server/om/sbin/pack
./queryManager.sh
```

  After the query script is executed, the following information is displayed:

```
Version          Package                          Cputype
***              FusionInsight_Manager_***        x86_64
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.4 Managing an MRS Cluster

8.4.1 Viewing Basic Information About an MRS Cluster

You can monitor and manage the clusters you have created. On the **Active Clusters** page, click the name of a cluster to open the cluster details page. On the displayed page, check the cluster's basic, network, and node information.

Viewing Basic Information About an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Dashboard** to view the basic information of the cluster.

Table 8-5 MRS cluster basic information

Parameter	Description
Cluster Name	The name of a cluster. Configure this parameter when creating a cluster. Click  to change the cluster name. Only the cluster name displayed on the MRS management console is changed, while the cluster name on MRS Manager is not changed synchronously.
Cluster Status	The cluster status. For details, see **Table 8-9**.
Cluster Version	MRS version information.
Cluster Type	There are three types of clusters: • **Analysis Cluster**: is used for offline data analysis and provides Hadoop components. • **Streaming Cluster**: is used for streaming tasks and provides stream processing components. • **Hybrid Cluster**: is used for both offline data analysis and streaming processing and provides Hadoop components and streaming processing components. • **Custom**: An MRS cluster with all custom components. MRS 3.*x* or later supports this type.
Cluster ID	Unique identifier of a cluster, which is automatically assigned when a cluster is created.
Created	Time when a cluster is created.
AZ	Availability zone (AZ) in the region of a cluster, which is set when a cluster is created.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Kerberos Authentication	Whether to enable Kerberos authentication when logging in to Manager.
Enterprise Project	The enterprise project to which a cluster belongs. Only on the **Active Clusters** page, you can click the name of an enterprise project to go to its **Enterprise Project Management** page.

Table 8-6 MRS cluster network information

Parameter	Description
Default Subnet	The subnet selected during cluster creation. If the subnet IP addresses are insufficient, click **Change Subnet** to switch to another subnet in the same VPC of the current cluster to obtain more available subnet IP addresses. Changing a subnet does not affect the IP addresses and subnets of existing nodes. A subnet provides dedicated network resources that are isolated from other networks, improving network security.
VPC	VPC selected during cluster creation. A VPC is a secure, isolated, and logical network environment.
EIP	After binding an EIP to an MRS cluster, you can use the EIP to access the Manager web UI of the cluster. If you do not need the EIP, click **Unbind** to unbind the EIP from the cluster. The Manager will not be accessible through the unbound EIP. **After an EIP is bound to the MRS Manager of a cluster, do not unbind it. Otherwise, other users may be unable to access the MRS Manager of that cluster.**
Security Group	The security group name of the cluster.

Table 8-7 MRS cluster O&M management

Parameter	Description
MRS Manager	Manager entry. For details, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
IAM User Sync	IAM user information (including federated users) can be synchronized to an MRS cluster for cluster management. For details, see **Synchronizing IAM Users to MRS**. • The **Components**, **Tenants**, and **Backups & Restorations** tab pages on the cluster details page can be accessed only after IAM users have been synchronized. In MRS clusters earlier than 3.*x*, the **Components** tab page can be used without IAM user synchronization. • For a federated user, only the information of the user who logged in to the system can be synchronized.
Data Connection	Click **Manage** to view the data connection type associated with the cluster. For details, see **Creating a Data Connection**.
Agency	Click **Manage Agency** to bind or modify an agency for the cluster. An agency allows ECS or BMS to manage MRS resources. You can configure an agency of the ECS type to automatically obtain the AK/SK to access OBS. For details, see **Interconnecting an MRS Cluster with OBS Using an IAM Agency**. The **MRS_ECS_DEFAULT_AGENCY** agency has the OBSOperateAccess permission of OBS and the CESFullAccess (for users who have enabled fine-grained policies), CES Administrator, and KMS Administrator permissions in the region where the cluster is located.
OBS Permission Control	Click **Manage** and modify the mapping between MRS users and OBS permissions. For details, see **Configuring Fine- Grained OBS Access Permissions for MRS Cluster Users**.
Logging	Used to collect logs about cluster creation and scaling failures.
Secure Communications	Used to display the security authorization status. You can enable or disable security authorization. Disabling security authorization brings high risks. Exercise caution when performing this operation. For details, see **Configuring Secure Communication Authorization for an MRS Cluster**.

----End

8.4.2 Checking the Running Status of an MRS Cluster

MRS allows you to create multiple clusters. The cluster quantity is subject to that of ECSs. You can view the running status of all MRS clusters on the console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Checking the Running Status of an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Click **Active Clusters**.

By default, clusters in the cluster list are sorted by creation time. For details about the cluster list parameters, see **Table 8-8**.

- **Active Clusters**: contain all clusters except the clusters in the **Failed** and **Deleted** states.
- **Cluster History**: contains the clusters in the **Deleted** states. Only clusters deleted within the last six months are displayed. If you want to view clusters deleted six months ago, contact technical support.
- **Failed Tasks**: Click  to view the failed cluster creation tasks.

Table 8-8 Parameters in the active cluster list

Parameter	Description
Name/ID	Cluster name, which is set when a cluster is created. Unique identifier of a cluster, which is automatically assigned when a cluster is created. • : Change the cluster name. • : Copy the cluster ID.
Cluster Version	Cluster version.
Cluster Type	The type of a cluster you want to create.
Nodes	Number of nodes that can be deployed in a cluster. This parameter is set when a cluster is created.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Status	Cluster running state and state change information. For details, see **Table 8-9**. The cluster creation progress includes: • Verifying cluster parameters • Applying for cluster resources • Creating VMs • Initializing VMs • Installing MRS Manager • Deploying the cluster • Cluster installation failed The cluster scale-out progress includes: • Preparing for scale-out • Creating VMs • Initializing VMs • Adding nodes to the cluster • Scale-out failed The cluster scale-in progress includes: • Preparing for scale-in • Decommissioning instance • Deleting VMs • Deleting nodes from the cluster • Scale-in failed The system will display causes of cluster installation, scale-out, and scale-in failures. For details, see **Table 5-8**.
Created	The time when a cluster node is successfully created. This parameter is displayed only on the **Cluster History** page.
Deleted	Time when a cluster node stops and the cluster node begins to be deleted. This parameter is valid only for historical clusters displayed on the **Cluster History** page.
AZ	Availability zone (AZ) in the region of a cluster, which is set when a cluster is created.
Enterprise Project	Enterprise project to which a cluster belongs.

Table 8-9 MRS cluster running state

Status	Description
Starting	When a cluster is being created, it is in the **Starting** state.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Status	Description
Running	If a cluster is created successfully and functioning properly, it is in the **Running** state.
Scaling out	If the Master, Core, or Task node in a cluster is being added, the cluster is in the **Scaling out** state.
Scaling in	If a cluster node is being deleted, the cluster node is in the **Scaling in** state. This state shows when you scale in or elastically scale in a cluster node, decommission a node in a yearly/monthly cluster, change the OS, or reinstall the OS.
Abnormal	If some components in a cluster are abnormal, the cluster is in the **Abnormal** state.
Deleting	When you delete a cluster node, the cluster state changes to **Deleting**. This state is shown after you click **Delete** and confirm the deletion.
Restoring node...	If a faulty node in the cluster is being recovered, its state is **Restoring node...**.

----End

8.4.3 Starting and Stopping an MRS Cluster

You can stop an MRS cluster using the MRS console or Manager when it is no longer needed or has been fixed due to a fault. Once stopped, the cluster's components will no longer serve external systems.

You can also start a stopped cluster.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Starting and Stopping a Cluster on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, choose **Management Operations** > **Stop All Components** in the upper right corner.

After the cluster stops, the status of all components in the component list changes to **Stopped**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

WARNING

This operation may interrupt jobs, cause data loss, or stop the cluster, adversely affecting service processing.

Step 4 For a cluster where all components have been stopped, choose **Management Operations** > **Start All Components** in the upper right corner of the cluster details page, perform operations as prompted, and wait until the components start.

After the operation is successful, **Operating Status** of the corresponding components changes to **Started**.

----End

Starting and Stopping a Cluster Using FusionInsight Manager

- MRS 3.x and later versions:
  a. Log in to FusionInsight Manager of the MRS cluster.
     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
  b. Click ... next to the target cluster and select **Stop**. Enter the password of the cluster administrator. In the **Stop Cluster** dialog box that is displayed, click **OK**. Wait until the cluster is stopped.
- MRS 2.x and earlier versions:
  a. Log in to FusionInsight Manager of the MRS cluster.
     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
  b. Click **Services**.
  c. In the upper part of the service list, choose **More** > **Start Cluster** or **Stop Cluster** accordingly.

Helpful Links

- For details about how to start or stop an MRS cluster component, see **Starting and Stopping an MRS Cluster Component**.
- For details about how to check the MRS cluster status, see **Checking the Running Status of an MRS Cluster**.

8.4.4 Restarting an MRS Cluster

During MRS cluster running, restart the components in the cluster if you have modified component settings, encountered infrastructure resource faults, or detected service process errors.

Components within an MRS cluster support both the standard restart and rolling restart.

- Standard restart: concurrently restarts all components or instances in the cluster, which may interrupt services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Rolling restart: restarts required components or instances without interrupting services as much as possible. Compared with a standard restart, a rolling restart takes a longer time and may affect service throughput and performance. To minimize or eliminate the impact on services during a component restart, you can perform rolling restarts to restart components or instances in batches. For instances in active/standby mode, the standby instance is restarted first, followed by the active instance.

  Table 8-11 describes the impact on services when a rolling restart is performed on components.

WARNING

Restarting a cluster will stop the cluster components from providing services, which adversely affects the running of upper-layer applications or jobs. You are advised to perform rolling restarts during off-peak hours.

Notes and Constraints

- Not all components in an MRS cluster support rolling restarts. When you perform a rolling restart on a cluster, components that do not support rolling restarts will be restarted in standard restart mode. For details, see **Component Restart Reference Information**.
- For configurations that must take effect immediately, for example, server port configurations, a standard restart is generally recommended over a rolling restart.

Impact on the System

Table 8-11 describes the impact of the rolling restart on each component during the rolling restart of a cluster.

Restarting a Cluster on the Console

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters** and click a cluster name to go to the cluster details page.

Step 3 In the upper right corner of the page, choose **Management Operations** > **Perform Rolling Cluster Restart**.

Step 4 The **Rolling-restart Cluster** page is displayed. Select **Only restart instances whose configurations have expired** and click **OK** to perform rolling restart for the target cluster.

Step 5 After the rolling restart task is complete, click **Finish**.

----End

Restarting a Cluster on FusionInsight Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Dashboard**. In the upper right corner, click **More** > **Restart**.

- For MRS 3.3.0 or later, the **Cluster** > **Dashboard** page has been removed from FusionInsight Manager. You can choose **More** in the upper right corner of **Homepage** to access cluster maintenance and management functions.
- When restarting a cluster, you can choose between a standard restart or rolling restart. A rolling restart takes a longer time but has less impact on services.

Step 3 In the displayed dialog box, enter the password of the current login user and click **OK**.

Step 4 If you select the rolling restart, adjust related parameters based on the site requirements.

Table 8-10 Rolling restart parameters

Parameter	Example Value	Description
Restart only instances with expired configurations in the cluster	-	Whether to restart only the modified instances in a cluster.
Enable rack strategy	-	Whether to enable the concurrent rack rolling restart strategy. This parameter takes effect only for roles that meet the rack rolling restart strategy. (The roles support rack awareness, and instances of the roles belong to two or more racks.) This parameter is available only when a rolling restart is performed on HDFS or YARN.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Data Nodes to Be Batch Restarted	1	Number of instances that are restarted in each batch when the batch rolling restart strategy is used. The default value is **1**. ● This parameter is valid only when the batch rolling restart strategy is used and the instance type is DataNode. ● This parameter is invalid when the rack strategy is enabled. In this case, the cluster uses the maximum number of instances (**20** by default) configured in the rack strategy as the maximum number of instances that are concurrently restarted in a rack. ● This parameter is configurable only when a rolling restart is performed on HDFS, HBase, YARN, Kafka, Storm, or Flume. ● This parameter for the RegionServer of HBase cannot be manually configured. Instead, it is automatically adjusted based on the number of RegionServer nodes. Specifically, if the number of RegionServer nodes is less than 30, the parameter value is **1**. If the number is greater than or equal to 30 and less than 300, the parameter value is **2**. If the number is greater than or equal to 300, the parameter value is 1% of the number (rounded-down).
Batch Interval	0	Interval between two batches of instances to be roll-restarted. The default value is **0**.
Decommissioning Timeout Interval	1800	Decommissioning interval for role instances during a rolling restart. The default value is **1800s**. Some roles (such as HiveServer and JDBCServer) stop providing services before the rolling restart. Stopped instances cannot be connected to new clients. Existing connections will be completed after a period of time. An appropriate timeout interval can ensure service continuity. This parameter is configurable only when a rolling restart is performed on Hive, Spark, or Spark2x.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Batch Fault Tolerance Threshold	0	Tolerance times when the rolling restart of instances fails to be batch executed. The default value is **0**, which indicates that the rolling restart task ends after any batch of instances fails to restart.

WARNING

Advanced parameters, such as **Data Nodes to Be Batch Restarted**, **Batch Interval**, and **Batch Fault Tolerance Threshold**, should be properly configured based on site requirements. Otherwise, services may be interrupted or cluster performance may be severely affected.

Example:

- If **Data Nodes to Be Batch Restarted** is set to an unnecessarily large value, a large number of instances are restarted concurrently. As a result, services are interrupted or cluster performance is severely affected due to too few working instances.
- If **Batch Fault Tolerance Threshold** is set to an unnecessarily large value, services will be interrupted because a next batch of instances will be restarted after a batch of instances fails to restart.

Step 5 Click **OK**.

----End

Reference

Table 8-11 describes the impact on the system during the rolling restart of components and instances.

Table 8-11 Impact on the system

Component	Service Interruption	Impact on System
ClickHouse	During the rolling restart, if the submitted workloads can be complete within the timeout period (30 minutes by default), there is no impact.	Nodes undergoing a rolling restart reject all new requests. This affects single-replica services, the ON CLUSTER operation, and workloads dependent on the instances being rolling-restarted. If a request that is being executed is not complete within the timeout period (30 minutes by default), the request fails.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
DBService	All services are normal during the rolling restart.	During the rolling restart, alarms indicating a heartbeat interruption between the active and standby DBService nodes may be reported.
Doris	Doris services will not be interrupted during the rolling restart only when the following conditions are met: • ELB or DBalancer is used to connect to Doris for job submission. • Tables being read or written must have multiple replicas (three replicas are recommended). • If the submitted workloads are complete within the timeout period (30 minutes by default), the workloads are not affected. Otherwise, the workloads will fail.	During the rolling restart, the total resources decrease, affecting the maximum memory and CPU resources that can be used by jobs. In extreme cases, the jobs may fail due to insufficient resources. If a job times out (30 minutes by default), retry the job.
Flink	All services are normal during the rolling restart.	The FlinkServer UI cannot be accessed during the rolling restart.
Flume	To prevent service interruptions and data loss, the following conditions must be met: • Persistent cache, for example File Channel, is used. • Flume cascading is configured. • Flume client sink supports fault failover or load balancing.	• Data loss may occur if channels do not have persistent cache. • The performance deteriorates for a short period of time during the fault failover.
Guardian	All services are normal during the rolling restart.	None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
HBase	HBase read and write services are normal during the rolling restart.	• Real-time read and write performance may deteriorate during RegionServer rolling restart. • During the HMaster rolling restart, services other than the real-time read and write services (excluding BulkLoad) are affected. Creating a table (**create**) Creating a namespace (**create_namespace**) Disabling a table (**disable**, **disable_all**) Re-creating a table (**truncate** and **truncate_preserve**) Moving a region (**move**) Taking a region offline (**unassign**) Merging regions (**merge_region**) Splitting a region (**split**) Enabling balance (**balance_switch**) DR operations (**add_peer**, **remove_peer**, **enable_table_replication**, **disable_peer**, **show_peer_tableCFs**, **set_peer_tableCFs**, **enable_peer**, **disable_table_replication**, **set_clusterState_active**, and **set_clusterState_standby**) Querying the cluster status (**status**)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
HDFS	• An active/standby switchover is triggered for NameNodes. During the switchover, there is temporarily no active NameNode. As a result, the HDFS unavailable alarm may be reported, and running read/write tasks may encounter errors. However, HDFS services are not interrupted. • If a rolling restart is being performed on DataNodes, errors may be reported for some data that is being read or written. Read/write speed is affected when the client retries. • During the rolling restart of ZKFC, an active/standby NameNode switchover occurs. • During the rolling restart of HDFS, enabling the rack policy may affect service read and write operations.	If a third-party client is used, the reliability of the third-party client during the rolling restart cannot be guaranteed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
HetuEngine	• There are at least two HSFabric instances, and at least two HSFabric instances are used for interconnection. Cross-domain services are not interrupted during the rolling restart. • HetuEngine services are not interrupted during the rolling restart of HSBroker, HSConsole, and QAS. • The rolling restart of HetuEngine compute instances can be performed only when there are at least two Coordinator and Worker nodes, respectively. The rolling restart of compute instances does not interrupt services.	• HSFabric nodes undergoing the rolling restart reject all new requests. SQL requests that are being executed will fail if they are not complete within the timeout period (30 minutes by default). • During the rolling restart, you cannot perform O&M operations on the HSConsole page. • During the rolling restart of HetuEngine compute instances, the performance decreases by approximately 20%. If the memory consumed by SQL execution exceeds 80% of **query_max_total_memory**, SQL tasks fail.
Hive	During the rolling restart, services with execution time longer than the decommissioning timeout period may fail.	• If the execution time of an existing task exceeds the timeout interval of the rolling restart, the task may fail during the restart. You can retry the task if it fails. • Nodes undergoing a rolling restart do not receive new requests. The number of requests processed by other instance nodes increases, and these nodes occupy more resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
Kafka	During the rolling restart, the read and write of Kafka topics with multiple replicas are normal, but operations on Kafka topics with only a single replica are interrupted.	• Topics and partitions cannot be added, deleted, or modified. • When **acks** is set to **1** or **0** in Producer, the next Broker will be forcibly restarted if the replica data is not synchronized within 30 minutes during the rolling restart. For a dual-replica Partition whose replicas are in the two Brokers, when **unclean.leader.election.enable** is set to **true** in the server configuration, data may be lost; when **unclean.leader.election.enable** is set to **false**, the Partition may not have a leader for a period of time until the Broker starts.
KrbServer	All services are normal during the rolling restart.	During the rolling restart, Kerberos authentication of a cluster may take a longer time.
LdapServer	All services are normal during the rolling restart.	During the rolling restart, Kerberos authentication of a cluster may take a longer time.
MapReduce	None	• You cannot view the task logs during the rolling restart. • In the scenario where only one JobHistoryServer instance is deployed, stopping JobHistoryServer will make all upper-layer components that depend on it become faulty.
Ranger	All services are normal during the rolling restart.	The RangerAdmin, RangerKMS, and PolicySync instances of Ranger are configured in active-active mode, and these instances can provide services in turn during the rolling restart. While UserSync supports only one-instance configuration and users cannot be synchronized during the restart. The user synchronization period is 5 minutes, and UserSync takes a short time to restart. Therefore, the UserSync restart has little impact on user synchronization.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
Spark	Except the listed items, other services are not affected.	• When performing a rolling restart on HBase, you cannot create or delete Spark on HBase tables in Spark. • When you perform a rolling restart on HBase, an active/standby switchover is triggered for HMaster. During the switchover, the Spark on HBase function is unavailable. • If you have used the advanced Kafka APIs, interruptions and data loss may occur when Spark reads data from or writes data to Kafka during the rolling restart. • During the rolling restart of ZooKeeper, spark-beeline fails to start. • During the rolling restart of YARN, Spark jobs may trigger retries at the task, stage, and application levels. • During the rolling restart of JDBCServer of Spark, long-running tasks will be terminated.
YARN	• An active/standby switchover is triggered for ResourceManager nodes during the rolling restart. Running tasks will encounter errors, but services are not interrupted. • During the rolling restart of NodeManager, containers submitted to the node may be retried on other nodes.	During the rolling restart of YARN, tasks running on YARN may experience exceptions due to excessive retries.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
ZooKeeper	ZooKeeper read and write operations are normal during the rolling restart.	• The rolling restart affects the ClickHouse service. During the instance restart, ClickHouse tables become read-only on each QuorumPeer instance for approximately 10 seconds. • The rolling restart causes ZooKeeper disconnection in Loader HA. Loader automatically retries for three times, each for 10 seconds. If ZooKeeper still cannot be connected, the active LoaderServer becomes the standby one and then a new active LoaderServer will be elected.

8.4.5 Exporting MRS Cluster Configuration Parameters

Administrators can export key server configuration parameters of MRS cluster components for quick check or backup.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Exporting Cluster Configurations Using the Console

This operation applies to MRS 2.*x* and earlier versions.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Choose **Configuration** > **Export Cluster Configuration** in the upper right corner of the page to export cluster configurations to the local PC.

----End

Exporting Cluster Configurations Using Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Dashboard**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 3.3.0 or later, the **Cluster** > **Dashboard** page has been removed from MRS Manager. You can choose **More** in the upper right corner of the **Homepage** to access cluster maintenance and management functions.
- For MRS clusters of version 2.*x* or earlier, click **Services** and choose **More** > **Export Cluster Configuration** to export the cluster configurations to a local directory.

Step 3 Choose **More** > **Export Configurations** to export the configurations of all services in the cluster in batches.

----End

8.4.6 Synchronizing the MRS Cluster Configuration

If a new configuration needs to be delivered to all services in the cluster, or **Configuration Status** of multiple services changes to **Expired** or **Failed** after a configuration is modified, the configuration parameters of these services are not synchronized and do not take effect. In this case, synchronize the configurations and restart related service instances for the cluster so that the new parameters take effect for all services.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Impact on the System

- After the cluster configurations are synchronized, you need to restart the services whose configurations have expired. During the restart, the corresponding services are unavailable.
- The instances whose configuration has expired are unavailable during restart.

Synchronizing Cluster Configurations Using the Console

This operation applies to MRS 2.*x* and earlier versions.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 choose **Configure** > **Synchronize Configuration** in the upper right corner.

Step 4 In the displayed dialog box, select "Restart services and instances whose configuration have expired" and click **OK** to restart the service whose configurations have expired.

When **Operation successful** is displayed, click **Finish**. The cluster is started.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Synchronizing Cluster Configurations Using Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Dashboard**.

- For MRS 3.3.0 or later, the **Cluster** > **Dashboard** page has been removed from MRS Manager. You can choose **More** in the upper right corner of the **Homepage** to access cluster maintenance and management functions.
- For MRS clusters of version 2.*x* or earlier, click **Services** and choose **More** > **Synchronize Configuration**.

Step 3 On this page, choose **More** > **Synchronize Configuration**.

Step 4 In the dialog box that is displayed, click **OK**.

Step 5 On the cluster dashboard page, choose **More** > **Restart Configuration-Expired Instances**.

Step 6 In the dialog box that is displayed, enter the password of the current login user and click **OK**.

Step 7 In the displayed dialog box, click **OK**.

You can click **View Instance** to open the list of all expired instances and confirm that the instances have been restarted.

----End

8.4.7 Deleting an MRS Cluster

To delete from a cluster, wait until data analysis and storage are finished, or if the cluster is not functioning properly and cannot provide services. If an MRS cluster fails to be deployed, the cluster is automatically deleted or unsubscribed.

If the current MRS cluster has been connected to OBS (decoupled storage and compute or cold and hot data separation scenarios), you need to manually delete service data on OBS after deleting a component or MRS cluster.

Deleting a MRS Cluster

Step 1 Log in to the MRS console.

Step 2 In the navigation pane on the left, choose **Active Clusters**.

Step 3 In the cluster list, locate the row containing the cluster to be deleted, and click **Delete** in the **Operation** column.

The cluster status changes from **Running** to **Deleting**, and finally to **Deleted**. You can view the deleted cluster in **Cluster History**.

----End

8.4.8 Changing the VPC Subnet of an MRS Cluster

A subnet is a unique CIDR block with a range of IP addresses in a VPC. You can divide a VPC into several subnets to help you plan IP addresses. All resources in a

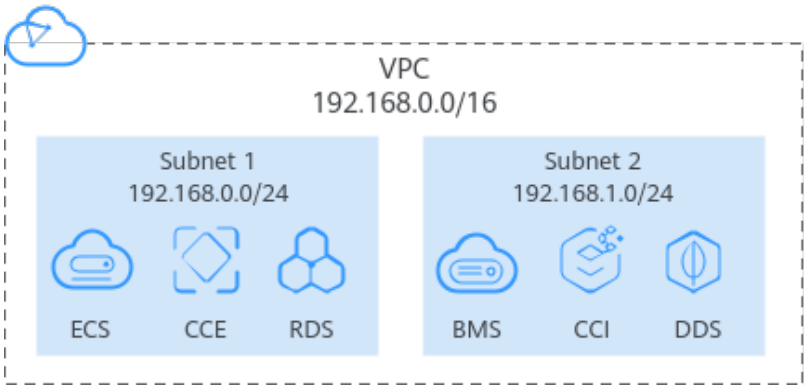
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

VPC must be deployed on subnets. Subnets in a VPC cannot overlap with each other.

- By default, all instances in different subnets of the same VPC can communicate with each other and the subnets can be located in different AZs.
- A cloud resource and its subnet can be in different AZs. For example, a cloud server in AZ 1 can use a subnet in AZ 3. If AZ 3 becomes faulty, the cloud server in AZ 1 can still use the subnet in AZ 3, and your services on the cloud server are not interrupted.

A network ACL protects an entire subnet. You can add inbound and outbound rules for a network ACL and associate a subnet with the network ACL. In this way, all instances in the subnet are protected by the network ACL. For details about how to configure outbound rules of a network ACL, see **How Do I Configure a Network ACL Outbound Rule?**.

Figure 8-1 Subnet



When creating an MRS cluster, you can manually select an existing subnet. After an MRS cluster is created, if IP addresses in the current subnet are insufficient, you can change the subnet to another one in the same VPC as the current cluster to increase available subnet IP addresses.

Changing a subnet will not have any impact on the IP addresses and subnets of existing nodes.

Changing a Subnet When No Network ACL Is Associated

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 In the **Network Information** area, click **Change Subnet** on the right of **Default Subnet**.

Step 4 Select the target subnet and click **OK**.

If no subnet is available, click **Create Subnet** to create a subnet first.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Changing a Subnet When a Network ACL Is Associated

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 In the **Network Information** area, view **VPC**.

Step 4 Log in to the VPC console. In the navigation pane on the left, choose **Virtual Private Cloud** and obtain the IPv4 CIDR block corresponding to the VPC obtained in **Step 3**.

Step 5 Choose **Access Control** > **Network ACLs** and click the name of the network ACL that is associated with the default and new subnets.

If both the default and new subnets are associated with a network ACL, add inbound rules to the network ACL by referring to **Step 6** to **Step 8**.

Step 6 On the **Inbound Rules** page, choose **More** > **Insert Rule Above** in the **Operation** column.

Step 7 Add a network ACL rule. Set **Action** to **Allow**, **Source** to the VPC IPv4 CIDR block obtained in **Step 4**, and retain the default values for other parameters.

Step 8 Click **OK**.

If you do not want to allow access from all IPv4 CIDR blocks of the VPC, add the IPv4 CIDR blocks of the default and new subnets by performing **Step 9** to **Step 13**. If you have added rules for the IPv4 CIDR block of the VPC, no action is required.

Step 9 Log in to the MRS console.

Step 10 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 11 In the **Network Information** area, click **Change Subnet** on the right of **Default Subnet**.

Step 12 Obtain the IPv4 CIDR blocks of the default and new subnets.

In this case, you do not need to click OK displayed in the Change Subnet dialog box. Otherwise, the default subnet will be updated to the new subnet, thereby making it difficult to query the IPv4 CIDR block of the default subnet. Exercise caution when performing this operation.

Step 13 Add the default subnet and the new subnet to the inbound rules of the network ACL by referring to **Step 5** to **Step 8**.

Step 14 Log in to the MRS console.

Step 15 Click the cluster name to go to its details page.

Step 16 In the **Network Information** area, click **Change Subnet** on the right of **Default Subnet**.

Step 17 Select the target subnet and click **OK**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Helpful Links

- Perform the following operations to configure an outbound rule for the network ACL:

How Do I Configure a Network ACL Outbound Rule?

- Method 1

  Allow all outbound traffic. This method ensures that clusters can be created and used properly.

- Method 2

  Allow the mandatory outbound rules that can ensure the successful creation of clusters. You are not advised to use this method because created clusters may not run properly due to absent outbound rules. If the preceding problem occurs, contact O&M personnel.

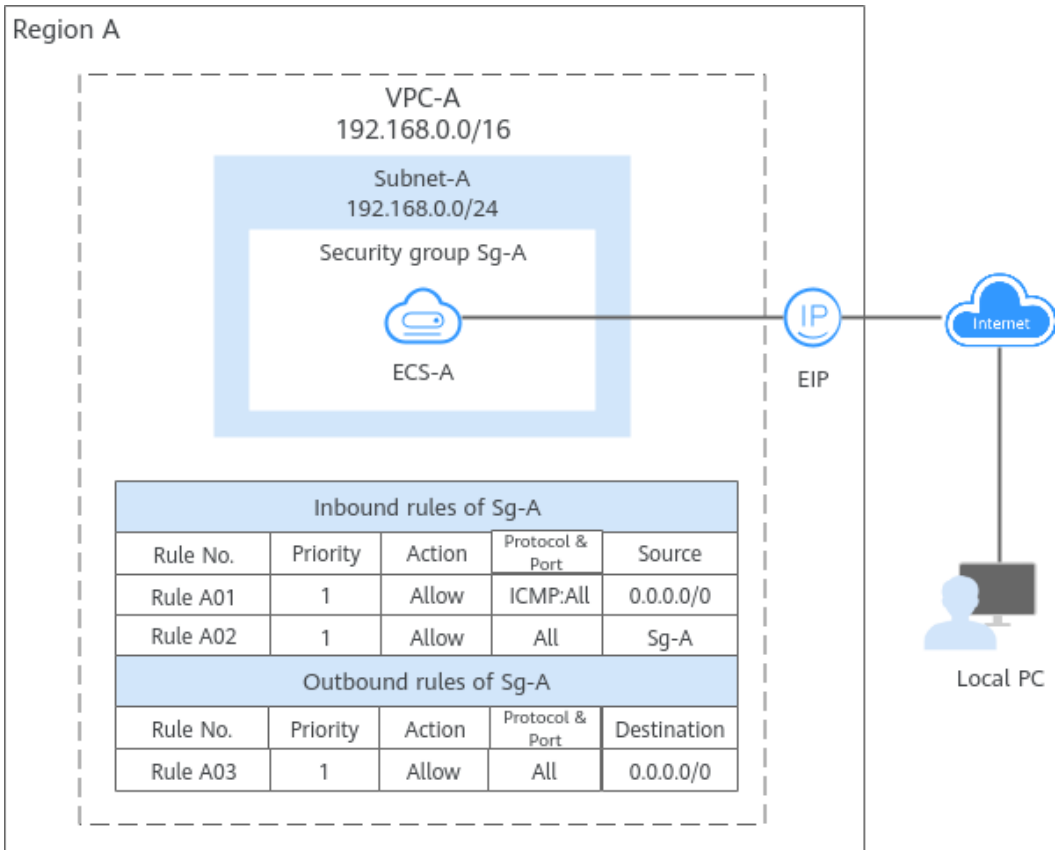
  Similar to the example provided in method 1, set **Action** to **Allow** and add the outbound rules whose destinations are the address with **Secure Communications** enabled, NTP server address, OBS server address, OpenStack address, and DNS server address, respectively.

8.4.9 Managing the Security Group of an MRS Cluster

A security group is a collection of access control rules for cloud resources, such as cloud servers, containers, and databases, that have the same security protection requirements and that are mutually trusted. A security group can have inbound and outbound rules. You need to specify the source, port, and protocol for each inbound rule and specify the destination, port, and protocol for each outbound rule to control the inbound and outbound traffic to and from the instances in the security group.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 8-2 Security group architecture



When creating an MRS cluster, you can enable the system to automatically create a security group or manually select an existing security group. Ensure that the security group used by the MRS cluster allows traffic from the **IP address range of the current user to reach TCP port 9022**. In this way, the user can access MRS Manager of the cluster.

After an MRS cluster is created, if you need to connect to nodes or big data components in the cluster, you need to add a security group rule to allow traffic on ports (for example, port 22) of the nodes or components.

Viewing the Security Group of an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 In the navigation pane, choose **Active Clusters**, select a running cluster, and click its name to switch to the cluster details page.

Step 3 In the network information area, view the security group of the MRS cluster.

Step 4 Open the VPC console.

Step 5 In the security group list, search for and locate the security group used by the MRS cluster. Click the security group name to go to its details page. You can then view details about the current security group, such as inbound and outbound rules.

Step 6 Click **Add Rule** to add an inbound or outbound rule to the security group.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If an EIP has been bound to MRS Manager of the current cluster, you can click **Add Security Group Rule** in the network information area of the cluster in **Step 3** to quickly add a security group rule to allow traffic on TCP port 9022.

----End

8.4.10 Modifying the OMS Service Configuration

Based on the security requirements of the user environment, you can modify the Kerberos and LDAP configurations in the OMS on FusionInsight Manager.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Impact on the System

After the OMS service configuration parameters are modified, the corresponding OMS module needs to be restarted. In this case, FusionInsight Manager cannot be used.

Modifying the OMS Service Configuration

Modifying the okerberos configuration

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **OMS**.

Step 3 Locate the row that contains okerberos and click **Modify Configuration**.

Step 4 Modify the parameters according to **Table 8-12**.

Table 8-12 okerberos parameters

Parameter	Example Value	Description
KDC Timeout (ms)	2500	Timeout duration for an application to connect to Kerberos, in milliseconds. The value must be an integer.
Max Retries	3	Maximum number of retries for an application to connect to Kerberos, in seconds. The value must be an integer.
LDAP Timeout (ms)	3000	Timeout duration for Kerberos to connect to LDAP, in milliseconds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
LDAP Search Timeout (ms)	3000	Timeout duration for Kerberos to query user information in LDAP, in milliseconds.
Kadmin Listening Port	21700	Port number of the Kadmin service.
KDC Listening Port	21701	Port number of the kinit service.
Kpasswd Listening Port	21702	Port number of the Kpasswd service.
Reset LDAP Account Password	-	Machine-machine users (**cn=krbadmin,ou=Users,dc=hadoop,dc=com** and **cn=krbkdc,ou=Users,dc=hadoop,dc=com**) used by Kerberos to access LDAP. If this parameter is selected, the passwords will be replaced by random passwords. **This option is available only in MRS 3.1.2 or later.**

Step 5 Click **OK**.

In the displayed dialog box, enter the password of the current login user and click **OK**. In the displayed confirmation dialog box, click **OK**.

Modifying the oldap configuration

Step 6 Locate the row that contains the oldap and click **Modify Configuration**.

Step 7 Modify the parameters according to **Table 8-13**.

Table 8-13 OLDAP parameters

Parameter	Example Value	Description
LDAP Listening Port	21750	Port number of the LDAP service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Reset LDAP Account Password	-	Machine-machine users (**cn=root,dc=hadoop,dc=com** and **cn=pg_search_dn,ou=Users,dc=hadoop,dc=com**) used by LDAP for data management, synchronization, and status check. If this parameter is selected, the passwords will be replaced by random passwords. **This option is available only in MRS 3.1.2 or later.**

Step 8 Click **OK**.

In the displayed dialog box, enter the password of the current login user and click **OK**. In the displayed confirmation dialog box, click **OK**.

Step 9 **If the LDAP account password is reset, restart ACS.**

1. Log in to the active management node as user **omm** and run the following command to update the domain configuration:
   ```
   sh ${BIGDATA_HOME}/om-server/om/sbin/restart-RealmConfig.sh
   ```
   The command is successfully executed if the following information is displayed:
   ```
   Modify realm successfully. Use the new password to log into FusionInsight again.
   ```
2. Run the following command to stop ACS:
   ```
   sh $CONTROLLER_HOME/sbin/acs_cmd.sh stop
   ```
3. Run the following command to start ACS:
   ```
   sh $CONTROLLER_HOME/sbin/acs_cmd.sh start
   ```

Restarting the cluster

Step 10 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 11 Restart the cluster. For details, see **Restarting an MRS Cluster**.

----End

8.4.11 Modifying MRS Manager Routing Table

When FusionInsight Manager is installed, two pieces of routing information are automatically created on the active management node. You can run the **ip rule list** command to view the routing information.

```
0:from all lookup local
32764:from all to 10.10.100.100 lookup ntp_rt   #NTP routing information created by FusionInsight Manager (this information is unavailable if no external NTP clock source is configured).
32765:from 192.168.0.117 lookup om_rt   #OM routing information created by the FusionInsight Manager.
32766:from all lookup main
32767:from all lookup default
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the routing information created by FusionInsight Manager conflicts with the routing information configured in the enterprise network planning, the cluster administrator can use **autoroute.sh** to disable or enable the routing information created by FusionInsight Manager.

Notes and Constraints

- When no external NTP server is configured, the system creates only OM routing information **om_rt**.
- This section applies only to MRS 3.*x* or later.

Impact on the System

After the routing information created by MRS Manager is disabled and before new routing information is set, MRS Manager cannot be accessed, but the clusters are running properly.

Prerequisites

You have obtained information about the route to be created from the cluster administrator.

Disabling System-Created Routing Information

Step 1 Log in to the active management node of the cluster as the **omm** user. Run the following commands to disable the routing information created by the system:

cd ${BIGDATA_HOME}/om-server/om/sbin

./autoroute.sh disable

```
Deactivating Route.
Route operation (disable) successful.
```

Step 2 View the execution result.

ip rule list

```
0:from all lookup local
32766:from all lookup main
32767:from all lookup default
```

Step 3 Run the following command and enter the password of user **root** to switch to user **root**:

su - root

Step 4 Run the following commands to manually create the routing information about the WS floating IP address:

ip route add *Network segment of the WS floating IP address/Subnet mask of the WS floating IP address* **scope link src** *WS floating IP address* **dev** *NIC of the WS floating IP address* **table om_rt**

ip route add default via *Gateway of the WS floating IP address* **dev** *NIC of the WS floating IP address* **table om_rt**

ip rule add from *WS floating IP address* **table om_rt**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Example:

ip route add 192.168.0.0/255.255.255.0 scope link src 192.168.0.117 dev eth0:ws table om_rt

ip route add default via 192.168.0.254 dev eth0:ws table om_rt

ip rule add from 192.168.0.117 table om_rt

NOTE

If IPv6 addresses are used, run the **ip -6 route add** command.

Step 5 Run the following commands to manually create the NTP service routing information. Skip this step when no external NTP clock source is configured.

ip route add default via *IP gateway of the NTP service* **dev** *NIC of the local IP address* **table ntp_rt**

ip rule add to *ntpIP* **table ntp_rt**

NIC of the local IP address indicates the NIC that can communicate with the network segment where the NTP server is located.

Example:

ip route add default via 10.10.100.254 dev eth0 table ntp_rt

ip rule add to 10.10.100.100 table ntp_rt

Step 6 View the execution result.

In the following example, if the command output contains **om_rt** and **ntp_rt**, the operation is successful.

ip rule list

```
0:from all lookup local
32764:from all to 10.10.100.100 lookup ntp_rt  #This information is not displayed if no external NTP clock source is configured.
32765:from 192.168.0.117 lookup om_rt
32766:from all lookup main
32767:from all lookup default
```

----End

Enabling System-Created Routing Information

Step 1 Log in to the active management node as user **omm**.

Step 2 Run the following commands to enable the routing information created by the system:

cd ${BIGDATA_HOME}/om-server/om/sbin

./autoroute.sh enable

```
Activating Route.
Route operation (enable) successful.
```

Step 3 View the execution result.

In the following example, if the command output contains **om_rt** and **ntp_rt**, the operation is successful.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ip rule list

```
0:from all lookup local
32764:from all to 10.10.100.100 lookup ntp_rt  #This information is not displayed if no external NTP clock source is configured.
32765:from 192.168.0.117 lookup om_rt
32766:from all lookup main
32767:from all lookup default
```

----End

8.5 Managing MRS Cluster Components

8.5.1 Checking the Running Status of an MRS Cluster Component

After creating an MRS cluster, you can check the status of each service component and its role instance on the console or Manager. This will assist you in detecting any performance issues with the component.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Checking Status on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the MRS cluster details page, click **Components** to view the service status, health status, and configuration status.

Table 8-14 Component status description

Status Type	Status	Description
Operating Status	Started	The service is started.
	Stopped	The service is stopped.
	Failed to start	Failed to start the role instance.
	Failed to stop	Failed to stop the role instance.
	Unknown	Initial service status after the background system restarts.
Health Status	Good	Indicates that all role instances in the service are running properly.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Status Type	Status	Description
	Faulty	The running status of at least one role instance is **Faulty** or the status of the service on which the current service depends is abnormal. If the running status of a service is **Faulty**, an alarm is generated. Rectify the fault based on the alarm information.
	Unknown	All role instances in the service are in the **Unknown** state.
	Restoring	The background system is restarting the service.
	Partially Healthy	When a dependent service fails, external systems cannot access its related APIs. This affects upper-layer businesses that rely on the service. HBase, Hive, Spark, and Loader may be in the **Partially Healthy** state. ● If YARN is installed but is abnormal, HBase is in the **Partially Healthy** state. ● If HBase is installed but is abnormal, Hive, Spark, and Loader are in the **Partially Healthy** state.
Configuration Status	Synchronized	Indicates that the latest configuration takes effect.
	Configuration expired	The latest configuration does not take effect after the parameter modification. You need to restart related services.
	Configuration failed	If a communication or read/write exception occurs during parameter configuration, you can use the synchronization configuration function to rectify the fault.
	Configuring	Parameters are being configured.
	Unknown	Current configuration status cannot be obtained.

Step 4 Click a component name to go to the component details page and view the detailed running information about the component.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Click **Instances** to view the detailed running information about each role instance in the service.

- The list of role instances shows all the instances for each role in the cluster, including their running and configuration status, hosts, and IP addresses.
- You can click an instance name to go to the instance details page and view the basic information, configuration file, instance logs, and monitoring metric graphs of the instance.

Table 8-15 Instance status

Status Type	Status	Description
Running Status	Good	The instance is running properly.
	Bad	The instance cannot run properly.
	Decommissioned	The instance is out of service.
	Not started	The instance is stopped.
	Unknown	The initial status of the instance cannot be detected.
	Starting	The instance is being started.
	Stopping	The instance is being stopped.
	Restoring	An exception may occur in the instance and the instance is being automatically rectified.
	Decommissioning	The instance is being decommissioned.
	Recommissioning	The instance is being recommissioned.
	Failed to start	The instance fails to be started.
	Failed to stop	The instance fails to be stopped.
Configuration Status	Synchronized	The latest configuration takes effect.
	Configuration expired	The latest configuration does not take effect after the parameter modification. You need to restart related services.
	Configuration failed	If a communication or read/write exception occurs during parameter configuration, you can use the synchronization configuration function to rectify the fault.
	Configuring	Parameters are being configured.
	Unknown	Current configuration status cannot be obtained.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

----End

Checking Status on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**. The component management page is displayed.

The service list displays the running status, configuration status, role type, and number of instances of each component.

On MRS Manager of MRS 2.*x* or earlier, click **Services** to open the component management page.

Table 8-16 Manager component status description

Status Type	Status	Description
Running Status	Normal	The component is running properly.
	Faulty	The component cannot work properly.
	Partially Healthy	Some enhanced functions of the component cannot work properly.
	Not started	The component is stopped.
	Unknown	The initial status of the component cannot be detected.
	Starting	The component is being started.
	Stopping	The component is being stopped.
	Failed to start	The component fails to be started.
	Failed to stop	The component fails to be stopped.
Configuration Status	Synchronized	The latest configuration takes effect.
	• Expired (Manager 2.*x* or earlier) • Expired (Manager 3.*x* and later)	The latest configuration does not take effect after the parameter modification. You need to restart related services.
	Failed	If a communication or read/write exception occurs during parameter configuration, you can use the synchronization configuration function to rectify the fault.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Status Type	Status	Description
	• Configuring (Manager 2.*x* or earlier) • Synchronizing (Manager 3.*x* and later)	Parameters are being configured.
	Unknown	Current configuration status cannot be obtained.

Step 3 Click a component name to view its details.

Step 4 Click **Instances** to view the detailed running information about each role instance in the service.

- The list of role instances shows all the instances for each role in the cluster, including their running and configuration status, hosts, and IP addresses.
- You can click an instance name to go to the instance details page and view the basic information, configuration file, instance logs, and monitoring metric graphs of the instance.

Table 8-17 Manager instance status description (3.x and later versions)

Status Type	Status	Description
Running Status	Normal	The instance is running properly.
	Faulty	The instance cannot run properly.
	Decommissioned	The instance is out of service.
	Not started	The instance is stopped.
	Unknown	The initial status of the instance cannot be detected.
	Starting	The instance is being started.
	Stopping	The instance is being stopped.
	Restoring	An exception may occur in the instance and the instance is being automatically rectified.
	Decommissioning	The instance is being decommissioned.
	Recommissioning	The instance is being recommissioned.
	Failed to start	The instance fails to be started.
	Failed to stop	The instance fails to be stopped.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Status Type	Status	Description
Configuration Status	Synchronized	The latest configuration takes effect.
	Expired	The latest configuration does not take effect after the parameter modification. You need to restart related services.
	Failed	If a communication or read/write exception occurs during parameter configuration, you can use the synchronization configuration function to rectify the fault.
	Synchronizing	Parameters are being configured.
	Unknown	Current configuration status cannot be obtained.

Table 8-18 Manager instance status description (2.x and earlier versions)

Status Type	Status	Description
Operating Status	Started	The role instance has been started.
	Stopped	The role instance has been stopped.
	Failed to start	Failed to start the role instance.
	Failed to stop	Failed to stop the role instance.
	Decommissioning	The role instance is being decommissioned.
	Decommissioned	The role instance has been decommissioned.
	Recommissioning	The role instance is being recommissioned.
	Unknown	Initial role instance status after the background system restarts.
Health Status	Good	The role instance is running properly.
	Restoring	The background system is restarting a role instance.
	Bad	The instance role is experiencing an abnormality, such as the inability to access a port due to a non-existent PID.
	Unknown	The host where a role instance resides does not connect to the background system.
	Partially Healthy	The role instance is partially running properly.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Status Type	Status	Description
Configuration Status	Synchronized	The latest configuration takes effect.
	Expired	The latest configuration does not take effect after the parameter modification. Related services need to be restarted.
	Failed	The communication is incorrect or data cannot be read or written during the parameter configuration. Click **Synchronize Configuration** to rectify the fault.
	Configuring	Parameters are being configured.
	Unknown	Current configuration status cannot be obtained.

----End

8.5.2 Starting and Stopping an MRS Cluster Component

You can stop a service component in the MRS cluster as required. After the component is stopped, it will no longer offer any services.

- Stop the services or abnormal services.
- Start the service in the **Stopped**, **Stop Failed**, or **Failed to Start** state to use the service again.
- Restart abnormal services or configure expired services to restore or enable the services.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Impact on the System

Some services within the system rely on one another. Therefore, when you start, stop, or restart a service, any services that depend on it will be affected.

- If a service is to be started, the lower-layer services dependent on it must be started first.
- If a service is stopped, the upper-layer services dependent on it are unavailable.
- If a service is restarted, the running upper-layer services dependent on it must be restarted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Starting and Stopping an MRS Cluster Component

You can start or stop a component on the management console or on FusionInsight Manager as required.

Starting and Stopping a Component on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the **Dashboard** page, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.

Step 4 On the cluster details page, click **Components**.

Step 5 Locate the row that contains the target component and click **Stop**.

Confirm the operation impact and click **OK**.

⚠ WARNING

After a component is stopped, it cannot provide services for external systems, which may affect the upper-layer services of the cluster. Ensure that the operation does not have any adverse impact before performing this operation.

After the operation is successful, **Operating Status** of the component changes to **Stopped**.

Step 6 To start a component, click **Start** in the **Operation** column of a stopped component and start the component as prompted.

After the operation is successful, **Operating Status** of the component changes to **Started**.

----End

Starting or Stopping a Component on FusionInsight Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

The service list contains all components installed in the cluster. When you select **Tile View**, the components are displayed as tiles. When you select **Table View**, the components are displayed in a table.

In this section, the **Table View** is used by default.

Step 3 Locate the target service and click **Stop Service** in the **Operation** column.

You can also click the name of the target service. On the displayed service details page, click **Stop Service** in the upper right corner.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Enter the password of the current administrator, confirm that the operation does not have any adverse impact, and click **OK**.

WARNING

After a component is stopped, it cannot provide services for external systems, which may affect the upper-layer services of the cluster. Ensure that the operation does not have any adverse impact before performing this operation.

After the operation is successful, **Running Status** of the component changes to **Not started**.

Step 5 To start a component, click **Start Service** in the **Operation** column of a stopped component and start the component as prompted.

After the operation is successful, **Running Status** of the component changes to **Normal**.

----End

8.5.3 Restarting an MRS Cluster Component

During MRS cluster running, restart the components in the cluster if you have modified component settings, encountered infrastructure resource faults, or detected service process errors.

Components within an MRS cluster support both the standard restart and rolling restart.

- Standard restart: concurrently restarts all components or instances in the cluster, which may interrupt services.
- Rolling restart: restarts required components or instances without interrupting services as much as possible. Compared with a standard restart, a rolling restart takes a longer time and may affect service throughput and performance. To minimize or eliminate the impact on services during a component restart, you can perform rolling restarts to restart components or instances in batches. For instances in active/standby mode, the standby instance is restarted first, followed by the active instance.

  Table 8-22 describes the impact on services when a rolling restart is performed on components.

WARNING

Restarting a cluster will stop the cluster components from providing services, which adversely affects the running of upper-layer applications or jobs. You are advised to perform rolling restarts during off-peak hours.

For details about whether services and instances in the current MRS cluster support rolling restart and the rolling restart parameters, see **Component Restart Reference Information**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Notes and Constraints

- Perform a rolling restart during off-peak hours.
  - If the service throughput of the Kafka service is high (over 100 MB/s) during a rolling restart, the restart will fail.
  - To avoid RegionServer restart failures caused by heavy loads during an HBase rolling restart, increase the number of handles if the requests per second of each RegionServer on the native interface exceed 10,000.
- Before restarting, check the current number of requests in HBase. If the number of requests on the native interface for each RegionServer is over 10,000, increase the number of handles to prevent overloading.
- Preferentially perform a rolling instance or service restart and select **Only restart instances whose configurations have expired**.

Impact on the System

- If the number of Core nodes in a cluster is less than six, services may be affected for a short period of time.
- **Table 8-22 describes the impact of a component rolling restart.**

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Restarting an MRS Cluster Component

Step 1 Access the MRS cluster component management page.

- Log in to the MRS console and click the cluster name to go to the cluster details page. Click **Components**.
- If you are using the Manager of MRS 3.*x* and later versions, log in to Manager and choose **Cluster** > **Services**.
- If you are using the Manager of MRS 2.*x* and earlier versions, log in to Manager and click **Services**.

Step 2 Click the name of the target component to go to the details page.

Step 3 On the service details page, expand the **More** drop-down list and select **Restart Service** or **Service Rolling Restart**.

Step 4 Enter the user password (required when you perform operations on Manager), confirm the operation impact, and click **OK** to restart the system.

If you select rolling restart, set parameters listed in **Table 8-19**. (Required parameters may vary by version, set parameters based on the actual GUI.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-19 Rolling restart configuration parameters

Parameter	Example Value	Description
Restart only instances with expired configurations	-	Whether to restart only the modified instances in a cluster. The name of this parameter may be different in other versions.
Enable rack strategy	-	Whether to enable the concurrent rack rolling restart strategy. This parameter takes effect only for roles that meet the rack rolling restart strategy. (The roles support rack awareness, and instances of the roles belong to two or more racks.) This parameter can be set only when a rolling restart is performed on HDFS or YARN.
Data Nodes to Be Batch Restarted	1	Number of instances that are restarted in each batch when the batch rolling restart strategy is used. The default value is **1**. ● This parameter is valid only when the batch rolling restart strategy is used and the instance type is DataNode. ● This parameter is invalid when the rack strategy is enabled. In this case, the cluster uses the maximum number of instances (20 by default) configured in the rack strategy as the maximum number of instances that are concurrently restarted in a rack. ● This parameter can be set only when a rolling restart is performed on some components, such as HDFS, HBase, YARN, Kafka, and Flume. The actual value displayed on the GUI prevails. ● The number of HBase RegionServer nodes that can be concurrently restarted during a rolling restart cannot be manually configured. It is automatically adjusted based on the total number of RegionServer nodes. The adjustment rules are as follows: If there are less than 30 nodes, one node is restarted at a time. If there are less than 300 nodes, two nodes are restarted at a time. If there are 300 or more nodes, 1% of the total node count is restarted at a time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Batch Interval	10	Interval (in seconds) between restarting two batches of instances during a rolling restart. The default value is **0**. Setting the batch interval parameter can increase the stability of the big data component process during the rolling restart. You are advised to set this parameter to a non-default value, for example, **10**.
Decommissioning Timeout Interval	1800	How long (in seconds) a role instance waits to be terminated after being marked for decommissioning during a rolling restart. This parameter can be set only when a rolling restart is performed on Hive or Spark. Some roles (such as HiveServer and JDBCServer) stop providing services before the rolling restart. Stopped instances cannot be connected to new clients. Existing connections will be completed after a period of time. An appropriate timeout interval can ensure service continuity.
Batch Fault Tolerance Threshold	0	Tolerance times when the rolling restart of instances fails to be batch executed. The default value is **0**, which indicates that the rolling restart task ends after any batch of instances fails to restart.

Step 5 After the restart or rolling restart of a component is successful, **Running Status** of the component is **Normal** and **Configure Status** is **Synchronized**.

----End

Component Restart Reference Information

Table 8-20 provides services and instances that support or do not support rolling restart in the MRS cluster.

Table 8-20 Services and instances that support or do not support rolling restart

Service	Instance	Rolling Restart
Alluxio	AlluxioJobMaster	Yes
	AlluxioMaster	
ClickHouse	ClickHouseServer	Yes
	ClickHouseBalancer	
CDL	CDLConnector	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Instance	Rolling Restart
	CDLService	
Flink	FlinkResource	No
	FlinkServer	
Flume	Flume	Yes
	MonitorServer	
Guardian	TokenServer	Yes
HBase	HMaster	Yes
	RegionServer	
	ThriftServer	
	RESTServer	
HetuEngine	HSBroker	Yes
	HSConsole	
	HSFabric	
	QAS	
HDFS	NameNode	Yes
	Zkfc	
	JournalNode	
	HttpFS	
	DataNode	
Hive	MetaStore	Yes
	WebHCat	
	HiveServer	
Hue	Hue	No
Impala	Impalad	No
	StateStore	
	Catalog	
IoTDB	IoTDBServer	Yes
Kafka	Broker	Yes
	KafkaUI	No
Kudu	KuduTserver	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Instance	Rolling Restart
	KuduMaster	
Loader	Sqoop	No
MapReduce	JobHistoryServer	Yes
Oozie	Oozie	No
Presto	Coordinator	Yes
	Worker	
Ranger	RangerAdmin	Yes
	UserSync	
	TagSync	
Spark	JobHistory	Yes
	JDBCServer	
	SparkResource	
Storm	Nimbus	Yes
	UI	
	Supervisor	
	Logviewer	
Tez	TezUI	No
YARN	ResourceManager	Yes
	NodeManager	
ZooKeeper	Quorumpeer	Yes

Table 8-21 lists the instance startup duration.

Table 8-21 Restart duration for reference

Service	Restart Duration	Startup Duration	Remarks
IoTDB	3min	IoTDBServer: 3 min	-
CDL	2min	• CDLConnector: 1 min • CDLService: 1 min	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Restart Duration	Startup Duration	Remarks
ClickHouse	4min	• ClickHouseServer: 2 min • ClickHouseBalancer: 2 min	-
HDFS	10min+x	• NameNode: 4 min + *x* • DataNode: 2 min • JournalNode: 2 min • Zkfc: 2 min	*x* indicates the NameNode metadata loading duration. It takes about 2 minutes to load 10,000,000 files. For example, *x* is 10 minutes for 50 million files. The startup duration fluctuates with reporting of DataNode data blocks.
YARN	5min+x	• ResourceManager: 3 min + *x* • NodeManager: 2 min	*x* indicates the time required for restoring ResourceManager reserved tasks. It takes about 1 minute to restore 10,000 reserved tasks.
MapReduce	2min+x	JobHistoryServer: 2 min + *x*	*x* indicates the scanning duration of historical tasks. It takes about 2.5 minutes to scan 100,000 tasks.
ZooKeeper	2min+x	quorumpeer: 2 min + *x*	*x* indicates the duration for loading znodes. It takes about 1 minute to load 1 million znodes.
Hive	3.5min	• HiveServer: 3 min • MetaStore: 1 min 30s • WebHcat: 1 min • Hive service: 3 min	-
Spark2x	5min	• JobHistory2x: 5 min • SparkResource2x: 5 min • JDBCServer2x: 5 min	-
Flink	4min	• FlinkResource: 1 min • FlinkServer: 3 min	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Restart Duration	Startup Duration	Remarks
Kafka	2min+x	• Broker: 1 min + *x* • Kafka UI: 5 min	*x* indicates the data restoration duration. It takes about 2 minutes to start 20,000 partitions for a single instance.
Storm	6min	• Nimbus: 3 min • UI: 1 min • Supervisor: 1 min • Logviewer: 1 min	-
Flume	3min	• Flume: 2 min • MonitorServer: 1 min	-
Doris	2 min	• FE: 1 min • BE: 1 min • DBroker: 1 min	-

Table 8-22 describes the impact on the system during the rolling restart of components and instances.

Table 8-22 Impact on the system

Component	Service Interruption	Impact on System
ClickHouse	During the rolling restart, if the submitted workloads can be complete within the timeout period (30 minutes by default), there is no impact.	Nodes undergoing a rolling restart reject all new requests. This affects single-replica services, the ON CLUSTER operation, and workloads dependent on the instances being rolling-restarted. If a request that is being executed is not complete within the timeout period (30 minutes by default), the request fails.
DBService	All services are normal during the rolling restart.	During the rolling restart, alarms indicating a heartbeat interruption between the active and standby DBService nodes may be reported.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
Doris	Doris services will not be interrupted during the rolling restart only when the following conditions are met: • ELB or DBalancer is used to connect to Doris for job submission. • Tables being read or written must have multiple replicas (three replicas are recommended). • If the submitted workloads are complete within the timeout period (30 minutes by default), the workloads are not affected. Otherwise, the workloads will fail.	During the rolling restart, the total resources decrease, affecting the maximum memory and CPU resources that can be used by jobs. In extreme cases, the jobs may fail due to insufficient resources. If a job times out (30 minutes by default), retry the job.
Flink	All services are normal during the rolling restart.	The FlinkServer UI cannot be accessed during the rolling restart.
Flume	To prevent service interruptions and data loss, the following conditions must be met: • Persistent cache, for example File Channel, is used. • Flume cascading is configured. • Flume client sink supports fault failover or load balancing.	• Data loss may occur if channels do not have persistent cache. • The performance deteriorates for a short period of time during the fault failover.
Guardian	All services are normal during the rolling restart.	None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
HBase	HBase read and write services are normal during the rolling restart.	• Real-time read and write performance may deteriorate during RegionServer rolling restart. • During the HMaster rolling restart, services other than the real-time read and write services (excluding BulkLoad) are affected. Creating a table (**create**) Creating a namespace (**create_namespace**) Disabling a table (**disable**, **disable_all**) Re-creating a table (**truncate** and **truncate_preserve**) Moving a region (**move**) Taking a region offline (**unassign**) Merging regions (**merge_region**) Splitting a region (**split**) Enabling balance (**balance_switch**) DR operations (**add_peer**, **remove_peer**, **enable_table_replication**, **disable_peer**, **show_peer_tableCFs**, **set_peer_tableCFs**, **enable_peer**, **disable_table_replication**, **set_clusterState_active**, and **set_clusterState_standby**) Querying the cluster status (**status**)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
HDFS	• An active/standby switchover is triggered for NameNodes. During the switchover, there is temporarily no active NameNode. As a result, the HDFS unavailable alarm may be reported, and running read/write tasks may encounter errors. However, HDFS services are not interrupted. • If a rolling restart is being performed on DataNodes, errors may be reported for some data that is being read or written. Read/write speed is affected when the client retries. • During the rolling restart of ZKFC, an active/standby NameNode switchover occurs. • During the rolling restart of HDFS, enabling the rack policy may affect service read and write operations.	If a third-party client is used, the reliability of the third-party client during the rolling restart cannot be guaranteed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
HetuEngine	• There are at least two HSFabric instances, and at least two HSFabric instances are used for interconnection. Cross-domain services are not interrupted during the rolling restart. • HetuEngine services are not interrupted during the rolling restart of HSBroker, HSConsole, and QAS. • The rolling restart of HetuEngine compute instances can be performed only when there are at least two Coordinator and Worker nodes, respectively. The rolling restart of compute instances does not interrupt services.	• HSFabric nodes undergoing the rolling restart reject all new requests. SQL requests that are being executed will fail if they are not complete within the timeout period (30 minutes by default). • During the rolling restart, you cannot perform O&M operations on the HSConsole page. • During the rolling restart of HetuEngine compute instances, the performance decreases by approximately 20%. If the memory consumed by SQL execution exceeds 80% of **query_max_total_memory**, SQL tasks fail.
Hive	During the rolling restart, services with execution time longer than the decommissioning timeout period may fail.	• If the execution time of an existing task exceeds the timeout interval of the rolling restart, the task may fail during the restart. You can retry the task if it fails. • Nodes undergoing a rolling restart do not receive new requests. The number of requests processed by other instance nodes increases, and these nodes occupy more resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
Kafka	During the rolling restart, the read and write of Kafka topics with multiple replicas are normal, but operations on Kafka topics with only a single replica are interrupted.	• Topics and partitions cannot be added, deleted, or modified. • When **acks** is set to **1** or **0** in Producer, the next Broker will be forcibly restarted if the replica data is not synchronized within 30 minutes during the rolling restart. For a dual-replica Partition whose replicas are in the two Brokers, when **unclean.leader.election.enable** is set to **true** in the server configuration, data may be lost; when **unclean.leader.election.enable** is set to **false**, the Partition may not have a leader for a period of time until the Broker starts.
KrbServer	All services are normal during the rolling restart.	During the rolling restart, Kerberos authentication of a cluster may take a longer time.
LdapServer	All services are normal during the rolling restart.	During the rolling restart, Kerberos authentication of a cluster may take a longer time.
MapReduce	None	• You cannot view the task logs during the rolling restart. • In the scenario where only one JobHistoryServer instance is deployed, stopping JobHistoryServer will make all upper-layer components that depend on it become faulty.
Ranger	All services are normal during the rolling restart.	The RangerAdmin, RangerKMS, and PolicySync instances of Ranger are configured in active-active mode, and these instances can provide services in turn during the rolling restart. While UserSync supports only one-instance configuration and users cannot be synchronized during the restart. The user synchronization period is 5 minutes, and UserSync takes a short time to restart. Therefore, the UserSync restart has little impact on user synchronization.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
Spark	Except the listed items, other services are not affected.	• When performing a rolling restart on HBase, you cannot create or delete Spark on HBase tables in Spark. • When you perform a rolling restart on HBase, an active/standby switchover is triggered for HMaster. During the switchover, the Spark on HBase function is unavailable. • If you have used the advanced Kafka APIs, interruptions and data loss may occur when Spark reads data from or writes data to Kafka during the rolling restart. • During the rolling restart of ZooKeeper, spark-beeline fails to start. • During the rolling restart of YARN, Spark jobs may trigger retries at the task, stage, and application levels. • During the rolling restart of JDBCServer of Spark, long-running tasks will be terminated.
YARN	• An active/standby switchover is triggered for ResourceManager nodes during the rolling restart. Running tasks will encounter errors, but services are not interrupted. • During the rolling restart of NodeManager, containers submitted to the node may be retried on other nodes.	During the rolling restart of YARN, tasks running on YARN may experience exceptions due to excessive retries.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Service Interruption	Impact on System
ZooKeeper	ZooKeeper read and write operations are normal during the rolling restart.	• The rolling restart affects the ClickHouse service. During the instance restart, ClickHouse tables become read-only on each QuorumPeer instance for approximately 10 seconds. • The rolling restart causes ZooKeeper disconnection in Loader HA. Loader automatically retries for three times, each for 10 seconds. If ZooKeeper still cannot be connected, the active LoaderServer becomes the standby one and then a new active LoaderServer will be elected.

Helpful Links

- For details about how to start or stop a cluster component, see **Starting and Stopping an MRS Cluster Component**.
- For details about how to restart an MRS cluster, see **Restarting an MRS Cluster**.
- For details about how to restart a role instance, see **Managing MRS Role Instances**.

8.5.4 Adding and Deleting an MRS Cluster Component

When creating an MRS cluster, you can select service components to be included in the cluster. After the cluster is created, administrators can manually add or delete components for MRS on the console.

Notes and Constraints

- Normal clusters: Only **custom clusters** of MRS 3.1.2 or later allow you to add or delete a component.
- LTS clusters: Only **custom clusters** of MRS 3.1.2-LTS.3 or a later LTS version allow you to add or delete a component.

Prerequisites

- You have configured permissions for the user group to which the IAM users belong.

  Adding or deleting a service in a cluster is a high-risk operation. Bind the MRS FullAccess, MRS Administrator, Server Administrator, Tenant Guest, MRS Administrator, or Tenant Administrator policy to the user group before you perform this operation.

  For details about the permissions, see **Synchronizing IAM Users to MRS**.
- **You have synchronized IAM users.** (On the **Dashboard** page, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You can add services only on nodes or node groups that are in normal state.
- If the service to be deleted has upper-layer dependencies, the service cannot be deleted. Only one service can be deleted at a time.

Adding an MRS Cluster Component

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, choose **Components** and click **Add Service**.

Step 4 In the service list, select the services to be added and click **Next**.

- When you add a service, the underlying services that it depends on are automatically selected. You can add multiple services at the same time.
- Components (MapReduce, YARN, and HDFS) in the Hadoop service cannot be added separately.
- If Hadoop is not installed in a cluster, after Hadoop is added, you need to refresh the console and synchronize IAM users again to submit jobs on the job management page.
- After the Spark2x/Spark component is added, if you need to perform operations on Spark SQL through the Hue Web UI, restart the Hue service first.

Step 5 On the **Topology Adjustment** page, select the nodes where the service is to be deployed. For details about the deployment scheme, see **Table 4-5**.

Step 6 Confirm the operation impact and click **OK**. After the service is added, you can view the added service on the **Components** page. **Table 8-23** describes the component status.

Table 8-23 Component status description

Status Type	Status	Description
Operating Status	Started	The service is started.
	Stopped	The service is stopped.
	Failed to start	Failed to start the role instance.
	Failed to stop	Failed to stop the role instance.
	Unknown	Initial service status after the background system restarts.
Health Status	Good	Indicates that all role instances in the service are running properly.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Status Type	Status	Description
	Faulty	The running status of at least one role instance is **Faulty** or the status of the service on which the current service depends is abnormal. If the running status of a service is **Faulty**, an alarm is generated. Rectify the fault based on the alarm information.
	Unknown	All role instances in the service are in the **Unknown** state.
	Restoring	The background system is restarting the service.
	Partially Healthy	When a dependent service fails, external systems cannot access its related APIs. This affects upper-layer businesses that rely on the service. HBase, Hive, Spark, and Loader may be in the **Partially Healthy** state. ● If YARN is installed but is abnormal, HBase is in the **Partially Healthy** state. ● If HBase is installed but is abnormal, Hive, Spark, and Loader are in the **Partially Healthy** state.
Configuration Status	Synchronized	Indicates that the latest configuration takes effect.
	Configuration expired	The latest configuration does not take effect after the parameter modification. You need to restart related services.
	Configuration failed	If a communication or read/write exception occurs during parameter configuration, you can use the synchronization configuration function to rectify the fault.
	Configuring	Parameters are being configured.
	Unknown	Current configuration status cannot be obtained.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Deleting an MRS Cluster Component

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click **Components**, locate the row that contains the target component, and click **Delete**.

In the displayed dialog box, click **Yes** to delete the service.

- If the service to be deleted has upper-layer dependencies, the service cannot be deleted. Only one service can be deleted at a time.
- All installed services except Hadoop (HDFS, YARN, MapReduce), Ranger, DBService, KrbServer, LdapServer, and meta services can be deleted.
- Before deleting a service, back up the service data to prevent data loss.

----End

Helpful Links

If the current cluster does not allow you to add or delete components, create an MRS cluster and select components that meet the requirements. For details, see **Buying MRS Clusters**.

8.5.5 Modifying the Configuration Parameters of an MRS Cluster Component

In a big data cluster, adjusting server parameters is a critical approach to optimizing cluster performance, stability, and resource utilization. When issues such as improper resource allocation or limited task parallelism arise, you need to promptly adjust relevant component parameters.

To meet actual business needs, the MRS console provides the function to check and modify component configuration parameters. You can quickly modify the parameters on the console.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Impact on the System

- After configuring the HBase, HDFS, Hive, Spark, YARN, and MapReduce service attributes, you need to download and install the client again or update the client configuration file.
- After configuring attributes of a service, restart the service to make the configurations take effect if the service status is **Expired**. The service is unavailable during the restart.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the service configuration parameters are modified and then take effect after a restart, you need to download and install the client again or download the configuration file to update the client.

Notes and Constraints

The parameters of DBService cannot be modified when there is only one DBService role instance in the cluster.

Modifying Parameters on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the MRS cluster details page, click **Components**.

Step 4 Select the target service from the service list.

Step 5 Click **Service Configuration**. Switch **Basic** to **All**. All configuration parameters of the service are displayed in the navigation tree. The service name and role names are displayed from upper to lower in the navigation tree.

NOTE

The parameters under the service node are service-level configuration parameters, and the parameters under the role node are role-level configuration parameters. For details, see **Modifying MRS Role Instance Parameters**.

Step 6 In the navigation tree, select a specified parameter and change its value. You can also enter the parameter name in the **Search** box to search for the parameter and view the result.

- If you are not sure about the location of a parameter, you can enter the parameter name in search box in the upper right corner. The system searches for the parameter in real time and displays the result.
- To cancel the change to a parameter value, click .
- Adjust the parameter values based on the suggestions in the parameter description to avoid adverse impact on cluster running and stability.
- Select a port number from the value range on the right. Ensure that all port numbers in the same service are within the value range and are unique. Otherwise, the service fails to start.

Step 7 Click **Save Configuration**, save the parameters as prompted, and restart the service.

After the component configuration parameters are saved, some configurations do not take effect immediately. In this case, the configuration status of the component instance changes to **Configuration expired**. You can wait until the service is idle or all related configurations of other components are modified, and then restart the component to make the configurations take effect.

In the component list, locate the row that contains the target component, click **Restart** in the **Operation** column. Restart the component as prompted. After the component successfully restarts, its **Configuration Status** changes to **Synchronized**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In versions earlier than MRS 3.*x*, to update the queue configuration of YARN without restarting the service, choose **More** > **Refresh Queue** on the **Service Status** tab page.

----End

Modifying Parameters on Manager

Step 1 Log in to Manager and choose **Cluster** > **Services**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Click the name of the service to be operated and click **Configurations**.

For clusters of MRS 2.*x* or earlier versions, log in to MRS Manager and click **Services**.

The **Basic Configurations** page is displayed by default. To modify more parameters, click **All Configurations**. The navigation tree displays all configuration parameters of the service. The level-1 nodes in the navigation tree are service names or role names. The parameter category is displayed after the level-1 node is expanded.

NOTE

The parameters under the service node are service-level configuration parameters, and the parameters under the role node are role-level configuration parameters. For details, see **Modifying MRS Role Instance Parameters**.

Step 3 In the navigation tree, select the specified parameter category and change the parameter values on the right.

- If you are not sure about the location of a parameter, you can enter the parameter name in search box in the upper right corner. The system searches for the parameter in real time and displays the result.
- To cancel the change to a parameter value, click .
- Adjust the parameter values based on the suggestions in the parameter description to avoid adverse impact on cluster running and stability.
- Select a port number from the value range on the right. Ensure that all port numbers in the same service are within the value range and are unique. Otherwise, the service fails to start.

Step 4 Click **Save**. In the confirmation dialog box, click **OK**.

Wait until the message "Operation succeeded" is displayed. Click **Finish**. The configuration is modified.

After the component configuration parameters are saved, some configurations do not take effect immediately. In this case, the configuration status of the component instance changes to **Expired**. You can wait until the service is idle or all related configurations of other components are modified, and then restart the component to make the configurations take effect.

In the component list, locate the row that contains the service whose configurations have expired, and choose **More** > **Restart Service** in the **Operation** column. Restart the service as prompted. After the service successfully restarts, the configuration status changes to **Synchronized**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- To update the queue configuration of the YARN service without restarting service, choose **More** > **Refresh Queue** to update the queue for the configuration to take effect.
- When configuring the **flume.config.file** parameter of Flume, you can upload and download files. After a configuration file is uploaded, the old file will be overwritten. If the configuration is not saved and the service is restarted, the new file configuration does not take effect. Save the configuration in time.
- If you need to restart the service for the configuration to take effect after modifying service configuration parameters, choose **More** > **Restart Service** in the upper right corner of the service page.
- If  is displayed before a parameter, the parameter takes effect dynamically. After the configuration is saved, the parameter value is automatically updated to the configuration file. (This function is available only in MRS 3.2.0 or later.)

----End

8.5.6 Viewing the Modified Component Configuration Parameters of an MRS Cluster

MRS allows you to view service configuration parameter changes in a cluster with only a few clicks, helping you quickly locate faults and improve configuration management efficiency.

The MRS Manager page displays non-initial default values for each service, non-unified values for identical role instances, historical records of cluster configuration changes, and expired configuration parameters in the cluster.

On the MRS Manager page, O&M personnel can view node configuration file content, including environment variables and role configurations, to quickly verify configuration accuracy or access hidden configuration items.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Viewing the Modified Component Parameters

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Configurations**.

Step 3 Select an operation page based on the scenario.

- View all non-default values.
  a. Click **All Non-default Values**. The system displays the parameters whose values are different from the default values configured for each service, role, or instance in the current cluster.

     You can click  next to a parameter value to quickly restore the value to the default one. You can click  to view the historical modification records of the parameter.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If there are a large number of parameters to configure, you can filter the parameters in the filter box in the upper right corner of the page or enter keywords in the search box.

b. To change the values of the parameters, change the values according to the parameter description and click **Save**. In the dialog box that is displayed, click **OK**.

- View all non-uniform values.
  a. Click **All Non-uniform Values**. The system displays parameters with different role, service, instance group, or instance configurations in the current cluster.

     You can click  next to a parameter value and view the differences in the dialog box that is displayed.

  b. To change the value of a parameter, click  to cancel the configuration difference or manually adjust the parameter value, click **OK**, and then click **Save**. In the dialog box that is displayed, click **OK**.
- View the expired configuration items.
  a. Click **Expired Configurations**. Expired configuration items in the current cluster are displayed.
  b. You can filter services using the service filter box in the upper part of the page to view expired configurations of different services. Alternatively, you can enter keywords in the search box.
  c. Expired configuration items do not take effect completely. Restart the services or instances whose configurations have expired in a timely manner.
- View historical configuration records.
  a. Click **Historical Configurations**. The historical configuration change records of the current cluster are displayed. You can view details about parameter value changes, including the service to which the parameter belongs, parameter values before and after the modification, and parameter files.
  b. To restore a configuration change, click **Restore Configuration** in the **Operation** column of the target record. In the dialog box that is displayed, click **OK**.

     If **the same configuration of the same service instance** is modified for multiple times, the configuration can only be restored to the latest modification.

Some configuration items take effect only after the corresponding services are restarted. After saving the configurations, promptly restart the services or instances with expired configurations.

----End

Viewing the Role Instance Configuration File

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **Cluster** > **Services**.

Step 3 On the page that is displayed, click **Instances**.

Step 4 Click the name of the target instance. In the **Configuration File** area on the **Instance Status** page, the configuration file list of the instance is displayed.

Step 5 Click the name of the configuration file to be viewed to view the parameter values in the configuration file.

To obtain the configuration file, download the configuration file to the local PC.

When a cluster node is faulty, the configuration file cannot be accessed. Rectify the fault before attempting to view the configuration file.

----End

8.5.7 Synchronizing MRS Component Configuration Parameters

If you find that the **status of some components or instances expires or fails**, you can synchronize the configuration to restore the configuration state. If all services in the cluster are in the **Failed** state, synchronize the cluster configuration with the background configuration.

- If all services in the cluster are in the **Configuration failed** state, synchronize the cluster configuration with the background configuration.
- If all services in the cluster are in the **Configuration failed** state, synchronize the service configuration with the background configuration.

Impact on the System

- After synchronizing cluster or service configurations, restart the services whose configurations have expired. These services are unavailable during the restart.
- After synchronizing the role instance configuration, restart the role instance whose configuration has expired. The role instance is unavailable during the restart.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Synchronizing Configurations on the Console

Synchronizing cluster configurations

Step 1 Log in to the MRS console.

Step 2 Click the target cluster name.

Step 3 On the cluster details page, choose **Configuration** > **Synchronize Configuration** in the upper right corner.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This operation applies only to MRS 2.*x* and earlier versions.

Step 4 In the displayed dialog box, select "Restart services and instances whose configuration have expired" and click **OK** to restart the service whose configurations have expired.

When **Operation successful** is displayed, click **Finish**. The cluster is started.

----End

Synchronizing component configurations

Step 1 Log in to the MRS console and click the name of the target cluster.

Step 2 On the MRS cluster details page, click **Components**.

Step 3 Select the target service from the service list.

Step 4 On the **Service Status** page, choose **More** > **Synchronize Configuration** and operate as prompted.

----End

Synchronizing role instance configurations

Step 1 Log in to the MRS console and click the name of the target cluster.

Step 2 On the MRS cluster details page, click **Components**.

Step 3 Select the target service from the service list.

Step 4 On the service details page, click **Instances**.

Step 5 Click the target role instance from the role instance list.

Step 6 In the upper part of the role instance status and metric information, choose **More** > **Synchronize Configuration** and synchronize the configuration as prompted.

----End

Synchronizing Configurations Using Manager

Synchronizing cluster configurations

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

- For MRS 3.*x*, choose **Cluster** > **Dashboard**.

  For MRS 3.3.0 or later, the **Cluster** > **Dashboard** page has been removed from MRS Manager. You can choose **More** in the upper right corner of the **Homepage** to access cluster maintenance and management functions.
- For MRS 2.*x* or earlier, click **Services**.

Step 2 On this page, choose **More** > **Synchronize Configuration**.

Step 3 Perform operations based on the cluster version.

- MRS 2.*x* and earlier versions: Perform identity authentication, select "Restart services and instances whose configuration have expired", and click **OK**. No further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- MRS 3.*x* and later versions: Click **OK**. After the synchronization is complete, perform the operations in **Step 4** to restart the instance whose configurations have expired.

Step 4 Restart the instance whose configurations have expired. (Perform this operation only for MRS 3.*x* and later versions.)

1. On Manager, choose **Cluster** > **Dashboard**.

   For MRS 3.3.0 or later, the **Cluster** > **Dashboard** page has been removed from MRS Manager. You can choose **More** in the upper right corner of the **Homepage** to access cluster maintenance and management functions.
2. Choose **More** > **Restart Configuration-Expired Instances**.
3. In the dialog box that is displayed, enter the password of the current login user and click **OK**.
4. In the displayed dialog box, click **OK**.

   Click **View Instance** to open the list of all expired instances and confirm that the instances have been restarted.

----End

Synchronizing component configurations

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

- For MRS 3.*x*, choose **Cluster** > **Services**.
- For MRS 2.*x* and earlier versions, click **Services**.

Step 2 Click the specified service name and choose **More** > **Synchronize Configuration**.

Step 3 Perform operations based on the cluster version.

- MRS 2.*x* and earlier versions: Perform identity authentication, select "Restart services and instances whose configuration have expired", and click **OK**. No further action is required.
- MRS 3.*x* and later versions: Click **OK**. After the synchronization is complete, perform the operations in **Step 4** to restart the instance whose configurations have expired.

Step 4 Restart the instance whose configurations have expired. (Perform this operation only for MRS 3.*x* and later versions.)

1. On Manager, choose **Cluster** > **Dashboard**.

   For MRS 3.3.0 or later, the **Cluster** > **Dashboard** page has been removed from MRS Manager. You can choose **More** in the upper right corner of the **Homepage** to access cluster maintenance and management functions.
2. Choose **More** > **Restart Configuration-Expired Instances**.
3. In the dialog box that is displayed, enter the password of the current login user and click **OK**.
4. In the displayed dialog box, click **OK**.

   Click **View Instance** to open the list of all expired instances and confirm that the instances have been restarted.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Synchronizing role instance configurations

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

- MRS 3.*x*: Choose **Cluster** > **Services**.
- For MRS 2.*x* and earlier versions, click **Services**.

Step 2 Click the specified service name and click **Instances**.

Step 3 Click the target role instance from the role instance list.

Step 4 On the role instance information page, choose **More** > **Synchronize Configuration**.

Step 5 Perform operations based on the cluster version.

- MRS 2.*x* and earlier versions: In the displayed dialog box, select "Restart services and instances whose configuration have expired" and click **OK**.
- MRS 3.*x* and later versions: Click **OK**. After the synchronization is complete, perform the operations in **Step 4** to restart the instance whose configurations have expired.

Step 6 Restart the instance whose configurations have expired. (Perform this operation only for MRS 3.*x* and later versions.)

1. Click **Instances**.
2. Select the instance to be restarted and choose **More** > **Restart Instance**.
3. In the dialog box that is displayed, enter the password of the current login user and click **OK**.
4. In the displayed dialog box, confirm the impact and click **OK**.

   Click **View Instance** to open the list of all expired instances and confirm that the instances have been restarted.

----End

8.5.8 Adding Custom MRS Component Parameters

The big data components in the cluster fully support all parameters from the open source community. You can modify configuration parameters on the MRS console pr MRS Manager for common scenarios, but some client components may not support all open source parameters.

To modify or add configuration parameters not displayed on the page, use the configuration item customization function. The new parameters will be saved in the component's configuration file and take effect after the component is restarted.

Prerequisites

- You have understood the meanings of parameters to be added, configuration files that have taken effect, and the impact on components.
- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Impact on the System

- After configuring properties of a service, restart the service if the service status is **Expired**. During the restart, the service cannot be accessed.
- After configuring the HBase, HDFS, Hive, Spark, YARN and MapReduce service properties, download and install the client again or download the configuration file to update the client.

Adding Custom Parameters on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the MRS cluster details page, click **Components**.

Step 4 Select the target service from the service list.

Step 5 Click **Service Configuration**.

Step 6 In the configuration type drop-down box on the right side, switch **Basic** to **All**.

Step 7 In the navigation tree on the left, locate a level-1 node and select **Customization**. The system displays the customized parameters of the current component.

- You can customize parameters for services and roles as required.
- Adding custom parameters for a single role instance is not supported.
- The configuration files that save the newly added customized parameters are displayed in the **Parameter File** column. Each configuration file may support open source parameters with the same name. The effective result is determined by the order in which the component loads the configuration files.

Step 8 Based on the configuration files and parameter functions, locate the row where a specified parameter resides, enter the parameter name supported by the component in the **Parameter** column and enter the parameter value in the **Value** column.

- You can click  or  to add or delete a custom parameter. You can delete a customized parameter only after you click for the first time.
- If you want to cancel the modification of a parameter value, click  to restore it.

Step 9 Click **Save Configuration** and operate as prompted.

----End

Adding Custom Parameters on Manager

For MRS 3.*x* and later versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

Step 3 Click the specified service name on the service management page.

Step 4 Choose **Configurations** > **All Configurations**.

Step 5 In the navigation tree on the left, locate a level-1 node and select **Customization**. The system displays the customized parameters of the current component.

- You can customize parameters for services and roles as required.
- Adding custom parameters for a single role instance is not supported.
- The configuration files that save the newly added customized parameters are displayed in the **Parameter File** column. Each configuration file may support open source parameters with the same name. The effective result is determined by the order in which the component loads the configuration files.

Step 6 Locate the row where a specified parameter resides, enter the parameter name supported by the component in the **Name** column and enter the parameter value in the **Value** column.

You can click **+** or **-** to add or delete a customized parameter.

Step 7 Click **Save**. In the displayed **Save Configuration** dialog box, confirm the modification and click **OK**. After the system displays "Operation succeeded", click **Finish**. The configuration is saved.

Restart the expired service or instance for the configuration to take effect.

If is displayed before a parameter, this parameter takes effect dynamically. After the configuration is saved, the parameter value is automatically updated to the configuration file. is supported only in MRS 3.2.0 or later.

----End

For MRS 2.*x* and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Services**.

Step 3 Select the target service from the service list.

Step 4 Click **Service Configuration**.

Step 5 Set **Type** to **All**.

Step 6 In the navigation tree, select **Customization**. The custom parameters of the current component are displayed on Manager.

- You can customize parameters for services and roles as required.
- Adding custom parameters for a single role instance is not supported.
- The configuration files that save the newly added customized parameters are displayed in the **Parameter File** column. Each configuration file may support

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

open source parameters with the same name. The effective result is determined by the order in which the component loads the configuration files.

Step 7 Based on the configuration files and parameter functions, locate the row where a specified parameter resides, enter the parameter name supported by the component in the **Name** column and enter the parameter value in the **Value** column.

- You can click  or  to add or delete a user-defined parameter. You can delete a custom parameter only after you click  for the first time.
- Click  to cancel the change to a parameter value.

Step 8 Click **Save Configuration** and select **Restart the affected services or instances**. Click **OK** to restart the services.

After the system displays "Operation succeeded", click **Finish**. The service is started.

----End

Task Example: Adding Custom Hive Parameters

Hive depends on HDFS. By default, Hive accesses the HDFS client. The configuration parameters to take effect are controlled by HDFS in a unified manner.

For example, the HDFS parameter **ipc.client.rpc.timeout** affects the RPC timeout period for all clients to connect to the HDFS server. If you need to modify the timeout period for Hive to connect to HDFS, you can use the configuration customization function. After this parameter is added to the **core-site.xml** file of Hive, this parameter can be identified by the Hive service and its configuration overwrites the parameter configuration in HDFS.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the MRS cluster details page, click **Components**.

Step 4 Choose **Hive** > **Service Configuration**.

Step 5 In the configuration type drop-down box on the right side, switch **Basic** to **All**.

Step 6 In the navigation tree on the left, select **Customization** for the Hive service. The system displays the customized service parameters supported by Hive.

Step 7 In **core-site.xml**, locate the row that contains the **core.site.customized.configs** parameter, enter **ipc.client.rpc.timeout** in the **Parameter** column, and enter a new value in the **Value** column, for example, **150000**. The unit is millisecond.

Step 8 Click **Save Configuration** and operate as prompted.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.5.9 Managing MRS Role Instances

You can start a role instance that is in the **Stopped**, **Failed to stop** or **Failed to start** status, stop an unused or abnormal role instance or restart an abnormal role instance to recover its functions.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Managing Instances on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Components**.

Step 4 Select the target service from the service list.

Step 5 Click the **Instances** tab.

Step 6 Select the check box on the left of the target role instance.

Step 7 Click **More**, select operations such as **Start Instance**, **Stop Instance**, **Restart Instance**, **Rolling-restart Instance**, or **Delete Instance** based on site requirements.

For details about instance support for rolling restart and the relevant parameters, see **Component Restart Reference Information**.

----End

Managing Instances on MRS Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Go to the service instance page.

- MRS 3.*x* and later versions: Choose **Cluster** > Services, click the name of the target service, and click **Instances**.
- MRS 2.*x* and earlier versions: Click **Services** and click **Instances**.

Step 3 Select the check box on the left of the target role instance.

Step 4 Choose **More** > **Start Instance**, **Stop Instance**, or **Restart Instance** as prompted.

For details about instance support for rolling restart and the relevant parameters, see **Component Restart Reference Information**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Exporting and Importing Role Instance Configurations

For clusters of MRS 3.*x* and later

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** and click the name of the service to be operated in the service view.

Step 3 Click **Instances** and click the instance to be operated.

Step 4 Click **Instance Configuration** and click **Export** to export the configuration parameter file to the local PC.

Step 5 On the **Instance Configurations** page, click **Import**, select the configuration parameter file of the instance, and import the file.

----End

For MRS clusters of 2.*x* and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Services**.

Step 3 Select a service.

Step 4 Select a role instance or click **Instances**.

Step 5 Select a role instance on a specified host.

Step 6 Click **Instance Configuration**.

Step 7 Click **Export Instance Configuration** to export the configuration data of a specified role instance, and choose a path for saving the configuration file.

Step 8 Click **Import Instance Configuration** to import the configuration data of the specified role instance.

Click **Save Configuration** and select **Restart the role instance**. Click **OK**.

After the system displays "Operation succeeded", click **Finish**. The role instance is started.

----End

8.5.10 Managing MRS Role Instance Groups

With MRS, you can easily manage instance groups by grouping multiple instances based on a specified principle, such as nodes with the same hardware configuration.

Any changes made to the configuration parameters of an instance group will apply to all instances within that group.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Notes and Constraints

This section applies to MRS 3.*x* or later.

Managing MRS Role Instance Groups

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 On Manager, choose **Cluster** > **Services**.

Step 3 Click the specified service name on the service management page.

Step 4 On the displayed page, click **Instance Groups**.

Click  and configure parameters as prompted.

Table 8-24 Instance group configuration parameters

Parameter	Description
Group Name	Instance group name. The name can contain only letters, digits, underscores (_), hyphens (-), and spaces. It must start with a letter, digit, underscore (_), or hyphen (-), cannot end with a space, and can contain a maximum of 99 characters.
Role	Role to which an instance group belongs.
Copy From	Instance group where the configuration parameter values are copied from. If the value is null, default values are used.
Description	Instance group description. It can contain only letters, digits, commas (,), periods (.), underscores (_), spaces, and line breaks, and can contain a maximum of 200 characters.

Each instance must belong to only one instance group. By default, when an instance is installed for the first time, it belongs to the instance group *Role name*-**DEFAULT**.

Step 5 Click **OK**.

Step 6 You can perform the following operations on an instance group as needed.

- Modifying an instance group: On the **Instance Group** page, locate the target instance group, click , and modify parameters as prompted. The default instance group cannot be modified.
- Deleting an instance group: Click **Instance Groups**. On the **Instance Groups** page, locate the row that contains the target instance group. Click . In the displayed dialog box, click **OK**. The default instance group cannot be deleted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can delete unnecessary or unused instance groups. Before deleting an instance group, migrate all instances in the group to other instance groups. The default instance group cannot be deleted.

- Viewing instance group information: On the **Instance Groups** page, select a role. On the **Basic** page, view all instances of the instance group.

  To move an instance from an instance group to another, perform the following operations:

  a. Select the instance to be moved and click **Move**.

  b. In the displayed dialog box, select the instance group to which the instance will be moved.

     During migration, the configuration of the target instance group is automatically inherited. If the instance configuration is modified before migration, the modified instance configuration will be applied.

  c. Click **OK**.

     Restart any service or instance with expired configurations to make the new settings take effect.

- Configuring instance group parameters: On the **Instance Groups** page, select a role in the navigation pane, click **Configuration**, adjust the configuration parameters, and click **Save**. The configuration will take effect on all instances in the instance group.

----End

8.5.11 Modifying MRS Role Instance Parameters

Configuration parameters of each role instance can be modified. In the scenario where instances are migrated to a new cluster or the service is redeployed, the cluster administrator can import or export all configuration data of a service to quickly copy configuration results.

Modifying configuration parameters, exporting instance configurations, or importing instance configurations will not impact other instances.

Impact on the System

- After modifying the configuration of a role instance, you need to restart the instance if the instance status is **Expired**. The role instance is unavailable during restart.
- You need to download and update the client configuration files after configuring HBase, HDFS, Hive, Spark, YARN, and MapReduce service properties.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Modifying Parameters on the Console

Step 1 Log in to the MRS console.

Step 2 Click the target cluster name.

Step 3 On the MRS cluster details page, click **Components**.

Step 4 Select the target service from the service list.

Step 5 Click the **Instances** tab.

Step 6 Click the target role instance from the role instance list.

Step 7 Click the **Instance Configuration** tab.

Step 8 Switch **Basic** to **All** from the drop-down list on the right of the page. All configuration parameters of the role instance are displayed in the navigation tree.

Step 9 In the navigation tree, select a specified parameter and change its value. You can also enter the parameter name in the **Search** box to search for the parameter and view the result.

If you want to cancel the modification of a parameter value, click  to restore it.

Step 10 Click **Save Configuration** and operate as prompted.

After the configuration parameters of a role instance are modified, you need to restart the instance if the instance status is **Configuration expired**. You can select the instances whose configurations have expired on the **Instances** tab page, and choose **More** > **Restart Instance**.

----End

Modifying Parameters on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

For MRS 2.*x* and earlier versions, click **Services**.

Step 3 Click the specified service name and click **Instances**.

Step 4 Click the specified instance and select **Instance Configuration**.

By default, **Basic Configuration** is displayed. To modify more parameters, click **All Configurations**. All parameter categories supported by the instance are displayed on the **All Configurations** page.

Step 5 In the navigation tree, select the specified parameter category and change the parameter values on the right.

Enter the name of a parameter in the upper right corner to search for it. Search results are displayed on Manager in real-time.

Step 6 Click **Save**. In the confirmation dialog box, click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For MRS 2.x and earlier versions, click **Save Configuration**, select **Restart the role instance**, and click **OK** to restart the role instance.

Wait until the message "Operation succeeded" is displayed. Click **Finish**. The configuration is modified.

- After the configuration parameters of a role instance are modified, you need to restart the instance if the instance status is **Expired**. You can select the instances whose configurations have expired on the **Instances** tab page, and choose **More** > **Restart Instance**.
- If is displayed before a parameter, this parameter takes effect dynamically. After the configuration is saved, the parameter value is automatically updated to the configuration file. is supported only in MRS 3.2.0 or later.

----End

8.5.12 Perform an Active/Standby Switchover for MRS Role Instances

In an MRS cluster, an active/standby switchover between the active and standby instance nodes is a critical mechanism for ensuring component high availability and data reliability.

For role instances that are deployed in active/standby mode in an MRS cluster (for example, NameNode instances of the HDFS component), if the active instance node is faulty and cannot provide services, the system automatically triggers an active/standby switchover, and the standby instance node automatically becomes the active one.

In case of excessively high resource workloads on the active node, software process faults, or cluster fault drills, the administrator can manually trigger an active/standby switchover.

For role instances that can be deployed in active/standby mode in an MRS cluster, see **Overview**.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

For MRS 2.*x* and earlier versions, click **Services**. The parameters vary depending on the version.

Step 3 Click the name of the component (for example, HDFS) where the role instance for which you want to perform an active/standby switchover is located.

On the **Instances** tab page, you can check the active/standby status of the role instances.

Step 4 On the service dashboard page, click **More** and select **Perform** *Role Instance* **Switchover**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, click **More** and select **Perform NameNode Switchover** for HDFS.

Step 5 In the displayed dialog box, enter the password of the current login user and click **OK**.

In the dialog box that is displayed, confirm the operation impact and click **OK** to perform an active/standby switchover for role instances.

When an active/standby switchover is performed for NameNodes on HDFS, you must specify the NameService.

Step 6 After the role instance switchover is successful, click **Instances**. You can find that the original active instance node becomes the standby one, the original standby instance becomes the active one, and **Running Status** of the instances is **Normal**.

----End

8.5.13 Decommissioning and Recommissioning an MRS Role Instance

MRS clusters can store data across different Core nodes. If a Core or Task node becomes faulty, the cluster status may be abnormal. To address this, you can decommission the specified role instance in MRS so that it no longer provides services.

After fault rectification, you can recommission the role instance.

The following role instances can be decommissioned or recommissioned:

- DataNode role instance on HDFS
- NodeManager role instance on YARN
- RegionServer role instance on HBase
- ClickHouseServer role instance of ClickHouse (supported in MRS 3.1.2 or later)
- IoTDBServer role instance of IoTDB
- Broker role instance of Kafka (supported in MRS 2.x and earlier versions and MRS 3.5.0-LTS and later versions)
- KuduTserver role instance of Kudu (supported in MRS 3.5.0-LTS or later)
- Doris BE instance (supported in MRS 3.3.0-LTS or later)

Notes and Constraints

- If a role instance is out of service, you must recommission the instance to start it before using it again.
- For details about ClickHouseServer instance decommissioning restrictions, see **Constraints on ClickHouseServer Scale-in**.
- DataNode recommissioning and decommissioning
  - By default, if the number of the DataNodes is less than or equal to that of HDFS replicas, decommissioning cannot be performed. If there are only three HDFS replicas and less than four DataNodes in the system, decommissioning cannot be carried out. In such a scenario, an error will be reported, and Manager will exit the decommissioning process 30 minutes after attempting to perform it.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - You can enable quick decommissioning before decommissioning DataNodes for clusters of MRS 3.3.0 or later. In this case, when the number of DataNodes meets the value of **dfs.namenode.decommission.force.replication.min**, the system decommissions the nodes and adds HDFS copies at the same time. **If data is written during quick decommissioning, data may be lost. Exercise caution when performing this operation.** The parameters related to quick decommissioning are listed as follows. You can search for and view the parameters on the HDFS configuration page on Manager.

    dfs.namenode.decommission.force.enabled: whether to enable quick decommissioning for DataNode. If this parameter is set to **true**, the function is enabled.

    dfs.namenode.decommission.force.replication.min: minimum number of available copies of a block required for DataNode quick decommissioning. The value ranges from 1 to 3.
  - During MapReduce task execution, files with 10 replicas are generated. Therefore, if the number of DataNode instances is less than 10, decommissioning cannot be performed.
  - If the number of DataNode racks (the number of racks is determined by the number of racks configured for each DataNode) is greater than 1 before the decommissioning, and after some DataNodes are decommissioned, that of the remaining DataNodes changes to 1, the decommissioning will fail. Therefore, before decommissioning DataNode instances, you need to evaluate the impact of decommissioning on the number of racks to adjust the DataNodes to be decommissioned.
  - If multiple DataNodes are decommissioned at the same time, and each of them stores a large volume of data, the DataNodes may fail to be decommissioned due to timeout. To avoid this problem, it is recommended that one DataNode be decommissioned each time and multiple decommissioning operations be performed.
- If the number of IoTDBServers is less than or equal to the number of region copies configured for the cluster (3 by default), decommissioning cannot be performed.
- Broker recommissioning and decommissioning (supported in MRS 3.5.0 and later versions):
  - During broker decommissioning, if the number of remaining brokers after decommissioning is less than the number of built-in topic replicas (3 by default) of the Kafka service, the broker cannot be decommissioned. If the instance is forcibly deleted, the service functions are unavailable.
  - If multiple brokers are decommissioned at the same time and the data volume of the topic partitions maintained by each broker is large, the brokers may fail to be decommissioned due to timeout. To avoid this problem, you are advised to decommission only one broker each time.
  - You can view the partition migration progress during broker decommissioning on the Kafka UI, and increase traffic limit to accelerate partition migration during off-peak hours. To avoid service interruption, you can reduce the partition migration traffic or cancel ongoing partition migration tasks.
  - Do not delete topics during broker decommissioning. Otherwise, there will be residual metadata of the migration tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - After adding a broker instance or recommissioning a broker, the system triggers partition balancing in 10 minutes. (You can use the **auto.reassign.check.interval.ms** parameter of the Kafka component on Manager to adjust the trigger time.)
- BE decommissioning
  - After Doris BE decommissioning, the number of remaining normal BE nodes must be no less than the replica count of any table. Otherwise, decommissioning will fail.
  - Before Doris BE decommissioning, the disk space of non-decommissioned BE nodes must be sufficient to store the data from all BE nodes to be decommissioned. In addition, about 10% of the storage space of each non-decommissioned BE node must remain reserved after decommissioning to ensure proper running of the remaining instances.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.
- If the DataNode is to be decommissioned, perform the following steps to perform a health check before decommissioning:
  a. Log in to the client installation node as a client user and switch to the client installation directory.
  b. For a security cluster, use user **hdfs** for permission authentication.
  ```
  source bigdata_env          #Configure client environment variables.
  kinit hdfs                  #Configure kinit authentication.
  Password for hdfs@HADOOP.COM:    #Enter the login password of user hdfs.
  ```
  c. Run the **hdfs fsck / -list-corruptfileblocks** command, and check the returned result.
     - If "has 0 CORRUPT files" is displayed, the health check is successful.
     - If the command output does not contain "has 0 CORRUPT files" and the name of the damaged file is returned, perform the following operations to delete the damaged file:

       hdfs dfs -rm *Name of the damaged file*

       WARNING

       Deleting a file is a high-risk operation. Ensure that the files are no longer needed before performing this operation.

Decommissioning an Instance on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 On the MRS cluster details page, click **Components**.

Step 4 Click a service in the service list.

Step 5 Click the **Instances** tab.

Step 6 Select an instance.

Step 7 Choose **More** > **Decommission** or **Recommission** to perform the corresponding operation.

During the instance decommissioning, if the service corresponding to the instance is restarted in the cluster using another browser, MRS displays a message indicating that the instance decommissioning is stopped, but the **Operating Status** of the instance is displayed as **Started**. In this case, the instance has been decommissioned on the background. You need to decommission the instance again to synchronize the operating status.

----End

Decommissioning an Instance on Manager

Step 1 Log in to MRS Manager and switch to the instance page of the target component.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

- MRS 3.*x* or later: Choose **Cluster** > **Services**, click the specified service name on the **Services** page, and click the **Instances** tab.
- MRS 2.*x* and earlier versions: Click **Services**, click the specified service name in the service list, and click **Instances**.

Step 2 Select the specified role instance to be decommissioned.

Step 3 Select **Decommission** or **Recommission** from the **More** drop-down list.

In the displayed dialog box, enter the password of the current login user and click **OK**.

Select the operation impact and click **OK** to perform the corresponding operation.

During instance decommissioning, if you restart the service of the instance in the cluster on another browser page, MRS may display a message indicating that the instance decommissioning is stopped, but the **Operating Status** of the instance is displayed as **Started**. In this case, the instance has been decommissioned on the background. You need to decommission the instance again to synchronize the operating status.

----End

8.5.14 Enabling and Disabling Ranger Authentication for an MRS Component

By default, the Ranger service is installed and Ranger authentication is enabled for a newly installed cluster with Kerberos authentication enabled. You can use the permission plug-in of the component to establish fine-grained security access policies for accessing component resources. If Ranger authentication is not required, the cluster administrator can manually disable Ranger authentication on the service page. After Ranger authentication is disabled, the system continues to

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

perform permission control based on the role model of Manager when accessing component resources.

In a cluster upgraded from an earlier version, Ranger authentication is not used by default when users access component resources. The cluster administrator can manually enable Ranger authentication after installing the Ranger service.

Notes and Constraints

- This section applies only to MRS 3.*x* or later.
- In a cluster in security mode, the following components support Ranger authentication: HDFS, YARN, Kafka, Hive, HBase, Storm, Impala, HetuEngine, CDL, and Spark/Spark2x.
- In a cluster in non-security mode, Ranger supports permission control on component resources based on OS users. The following components support Ranger authentication: HBase, HDFS, Hive, Spark/Spark2x, YARN, and HetuEngine (supported in MRS 3.3.0 and later).
- Once Ranger authentication is enabled, Ranger takes over all component authentication. Permissions set by the original authentication plug-in become invalid, except for the ACL rules of HDFS and YARN, which still take effect. Proceed with caution and ensure permissions are pre-configured in Ranger.
- Once Ranger authentication is disabled, the component's own permission plug-in takes over all component authentication. Permissions configured in Ranger become invalid. Proceed with caution and ensure permissions are pre-configured on MRS Manager.

Enabling Ranger Authentication

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

Step 3 Click the specified service name on the service management page.

Step 4 On the service details page, expand the **More** drop-down list and select **Enable Ranger**.

Step 5 In the displayed dialog box, enter the password of the current login user and click **OK**.

Step 6 In the service list, restart the service whose configuration has expired.

----End

Disabling Ranger Authentication

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click the specified service name on the service management page.

Step 4 On the service details page, expand the **More** drop-down list and select **Disable Ranger**.

Step 5 Enter the password of the current login user and click **OK**. In the displayed dialog box, click **OK**.

Step 6 In the service list, restart the service whose configuration has expired.

----End

8.5.15 Accessing Web UIs of Open-Source Components Managed in MRS Clusters

When using an MRS cluster, you can perform O&M and configuration operations on clusters, monitor cluster status, and manage users and role on MRS Manager. Some big data components in a cluster provide the web UIs, which allows you to quickly learn more about the component running status or configure services.

After logging in to MRS Manager, you can switch to the component web UI with a few clicks.

Table 8-25 describes the functions provided by web UIs of common components in an MRS cluster.

Table 8-25 Web UIs of common components in an MRS cluster

Component	Web UI Usage
Flink	Flink Web UI: a web-based, visualized development platform for Flink jobs. It enables you to quickly develop and manage Flink SQL and Flink JAR jobs.
HBase	HMaster Web UI: allows you to check basic information about an HBase cluster, including the tables, regions, and region servers in the cluster.
HDFS	NameNode Web UI: allows you to check the Hadoop cluster status, node information, and HDFS information.
HetuEngine	HSConsole Web UI: provides functions such as HetuEngine data source management, compute instance management, and automatic task viewing.
Hive	HiveServer2 Web UI: allows you to view Hive configurations, logs, metrics, and session information.
Hue	Hue Web UI: allows you to view content related to the Hadoop ecosystem, facilitating data analysis and cluster management. For example, you can view the HDFS file systems, Hive data tables, and job information within a cluster.
Impala	StateStore Web UI: displays health check and status information related to the state store, such as Impala memory usage and configuration settings.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Web UI Usage
Kudu	KuduMaster Web UI: allows you to monitor and manage Kudu clusters. For example, you can access related endpoints to view KuduMaster server logs on the web UI, helping you locate faults that occur during cluster running, such as data read/write exceptions and node connection faults.
Kafka	KafkaManager Web UI: allows you to manage and configure Kafka topics.
Loader	Loader Web UI: Loader is an ETL tool used to exchange data and files between MRS and external data sources such as relational databases, SFTP servers, and FTP servers.
MapReduce	JobHistoryServer Web UI: a tool for viewing the history of MapReduce jobs. It displays basic information such as the start time, end time, and input and output paths of jobs, helping you understand the overall execution period and data processing scope of jobs.
Presto	Coordinator Web UI: allows you to check and monitor Presto clusters, as well as manage and check running queries.
Ranger	Ranger Web UI: allows you to manage and configure fine-grained access control policies for components such as HDFS, HBase, Hive, and YARN within a cluster.
Spark	Spark Web UI: enables you to monitor the execution and resource consumption of Spark applications in real time. Using this web UI, developers and O&M personnel can quickly locate performance bottlenecks (such as data skew and memory overflow), monitor resource utilization, and optimize the execution efficiency of Spark applications.
Storm	Storm Web UI: allows you to monitor the topology running status, check node (Nimbus/Supervisor) information, and manages task allocation and resource usage.
Tez	Tez Web UI: displays the Tez task execution process on a GUI. You can view the task execution details on the web UI.
YARN	• ResourceManager Web UI: displays the total number of cluster resources, node status, application running status, and queue resource allocation. • TimelineServer Web UI: provides a graphical interface for querying and displaying historical events and metrics of YARN applications.

Prerequisites

In a cluster with Kerberos authentication enabled, user admin does not have the highest management permissions on components. To access the web UI of a component, you need to create a service user with the management permissions

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

on the component in advance. For details about how to create a user, see **Creating an MRS Cluster User**.

Accessing the Component Web UI

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Prepare a service user for accessing the component Web UI.

In a cluster with Kerberos authentication enabled, user admin does not have the highest management permissions on components. To access the web UI of a component, you need to create a service user with the management permissions on the component in advance. For details, see **Creating an MRS Cluster User**.

For example, if you need to view job details on the YARN Web UI, the current login user must have the permissions to log in to Manager and manage YARN jobs.

On Manager, create a human-machine user and associate the user with the **System_administrator** role.

Step 3 Log in to Manager as a user who has the permissions to manage the component Web UI. Choose **Cluster** > **Services**, and click the required service name. On the **Dashboard** page, click the web UI link in the **Basic Information** area.

You can click the web UI link on the dashboard page to access the relevant web UI. The following uses accessing the HDFS Web UI as an example. On the **Dashboard** page, click the web UI link **NameNode** (*Host name*,**Active**).

Step 4 On the component Web UI, perform other operations as required.

----End

8.6 Managing MRS Cluster Nodes

8.6.1 Checking the Running Status of an MRS Cluster Node

You can keep track of the status of each node in a running MRS cluster in real-time on the console or Manager. This helps you promptly detect any potential resource problems.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Checking Node Status on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Nodes**.

Step 4 Expand the node group name to view basic information about the nodes in a node group, such as their status, CPU usage, memory usage, and disk usage.

In the host list, click a host name to view detailed node information.

Table 8-26 MRS cluster node information

Parameter	Description
Node Group	Node group name.
Node Type	Node type corresponding to the node group. For details, see **MRS Cluster Node Types**. On the **Nodes** page, click ⌄ next to a node group name to unfold the nodes contained in the node group. Click a node name to remotely log in to the ECS using the password or key pair configured during cluster creation. For details about the parameters, see **Viewing MRS Cluster Component Monitoring Metrics**.
Node Count	Number of nodes in a node group.

----End

Checking Node Status on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts** to view the host list.

The host list shows all the hosts in the cluster along with their basic information. You can use this list to perform O&M operations on hosts, apply filters, and search for specific hosts. Additionally, you can click **Export All**, set **Save As** to **TXT** or **CSV**, and click **OK** to export information about all hosts. **Host View** is displayed by default. You can click **Role View** to switch the view type. You can also click the edit button to customize the content displayed in each view.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-27 View type description

View Type	Description
Host View	Displays the IP address, rack planning, running status, and hardware resource usage of each host.
Role view	Displays the roles that have been deployed on each host. If the role supports the active/standby mode, the role name is displayed in bold.

Table 8-28 Host running status description

Status	Description
Normal	Indicates that the host is in the normal state.
Faulty	Indicates that the host is abnormal.
Unknown	Indicates that the initial status of the host cannot be detected.
Isolated	Indicates that the host is isolated.
Suspended	Indicates that the host is stopped.

Step 3 In the host list, click the specified host name to view the host overview information.

The host details page contains the basic information area, disk status area, instance list area, and monitoring graphs.

Table 8-29 Host details

Item	Description
Basic Information	The basic information area contains the key information about the host, such as the management IP address, service IP address, host type, rack, firewall, number of CPU cores, and OS.
Disk	The disk status area contains all disk partitions configured for the cluster on the host and the usage of each disk partition.
Instances	The instance list displays all role instances installed on the host and the status of each role instance. You can click the log file next to a role instance name to view the log file content of the instance online.
Alarm and Event Record	The alarm and event history area displays the key alarms and events reported by the current host. The system can display a maximum of 20 historical records.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Item	Description
Chart	The monitoring chart area is displayed on the right of the host details page, and contains the key monitoring metrics of the host. You can choose ▼ > **Customize** in the upper right corner to customize the monitoring reports to be displayed in the chart area. Select a time range and choose ▼ > **Export** to export detailed monitoring metric data within the specified time range. You can click ⓘ next to the title of a monitoring indicator to open the description of the monitoring metric.

Step 4 View the host chart, process, and resource information.

Click the **Chart**, **Process**, and **Resource** to view the full monitoring chart information about the host.

- The **Chart** page displays all monitoring graphs of the host.
- The **Process** page shows information about the role processes of the deployed service instances on the current host, including their status, PID, and running time. You can easily access the log files of each process online.
- The **Resource** page displays the resource usage details of the deployed service instances on the current host, such as CPU, memory, disk, and port usage.

----End

8.6.2 Starting and Stopping All Roles on an MRS Cluster Node

When a host or node is faulty, you may need to stop all roles on the host on MRS to check the host. After the fault is rectified, start all roles on the host to recover host services.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Starting or Stopping Roles on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Nodes**.

Step 4 Unfold the node group information and select the check box of the target node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Choose **Node Operation** > **Start All Roles** or **Stop All Roles** to perform the required operation.

----End

Starting or Stopping Roles on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts** to go to the host list page.

For clusters of MRS 2.*x* or earlier, click **Hosts**.

Step 3 Select the check box of the target hosts.

Step 4 Select **Start All Instances** or **Stop All Instances** from the **More** drop-down list to start or stop all role instances.

----End

8.6.3 Isolating an MRS Cluster Node

If a host is found to be abnormal or faulty, affecting cluster performance or preventing services from being provided, you can temporarily exclude that host from the available nodes in the cluster. In this way, the client can access other available nodes. In scenarios where patches are to be installed in a cluster, you can also exclude a specified node from patch installation.

You can isolate a host manually on MRS based on the actual service requirements or O&M plan.

Notes and Constraints

Only non-management nodes can be isolated.

Impact on the System

- After a host is isolated, all role instances on the host will be stopped. You cannot start, stop, or configure the host and any instances on the host.
- After a host is isolated, statistics of the monitoring status and indicator data of the host hardware and instances cannot be collected or displayed.
- For some services, after a host is isolated, some instances on other nodes do not work, and the service configuration status may expire.
- Retain the default SSH port (22) of the target node. Otherwise, the task described in this section will fail.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Isolating a Host on the Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Nodes**.

Step 4 Unfold the node group information and select the check box of the target host.

Step 5 Choose **Node Operation** > **Isolate Host**.

Step 6 Confirm the information about the host to be isolated and click **OK**.

When **Operation successful** is displayed, click **Finish**. The host is isolated successfully, and the value of **Operating Status** becomes **Isolated**.

Step 7 Log in to the isolated host as user **root**, run the **pkill -9 -u omm** command to stop the processes of user **omm** on the node, and run the **ps -ef | grep 'container' | grep '${BIGDATA_HOME}' | awk '{print $2}' | xargs -I '{}' kill -9 '{}'** command to search for and stop the container process.

Step 8 Host isolation can be canceled, after which the host can be added back to the cluster. If the host exception or fault has been rectified, cancel the host's isolation status before reusing the host.

1. On the **Nodes** page, expand the node group information.
2. Select the check box of the host to be de-isolated and choose **Node Operation** > **Cancel Host Isolation**.
3. Confirm the information about the host for which the isolation is to be cancelled and click **OK**.

   When **Operation successful** is displayed, click **Finish**. The host is de-isolated, and the value of **Operating Status** becomes **Normal**.

----End

Isolating a Host on Manager

For MRS 3.*x* or later:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts**.

Step 3 Select the check box of the host to be isolated, choose **More** > **Isolate Host**, enter the password for identity authentication, and click **OK**.

Step 4 Confirm the information about the host to be isolated and click **OK**.

After the system displays a message indicating that the operation is successful, click **Finish**. If the host is isolated, its **Running Status** is displayed as **Isolated**.

Step 5 Log in to the isolated host as user **root** and run the **pkill -9 -u omm** command to stop the processes of user **omm** on the node. Then run the **ps -ef | grep**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

'container' | grep '${BIGDATA_HOME}' | awk '{print $2}' | xargs -I '{}' kill -9 '{}' command to find and stop the container process.

Step 6 Cancel the isolation status of the host before using the host if you have rectified the host exception or fault.

On the **Hosts** page, select the isolated host and choose **More** > **Cancel Isolation**.

When the isolation of a host is canceled, its role instances remain stopped by default. To start role instances on the host, select the target host on the **Hosts** page and choose **More** > **Start All Instances**.

----End

For MRS 2.*x* or earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts**.

Step 3 Select the check box of the host to be isolated, choose **More** > **Isolate Host**, enter the password for identity authentication, and click **OK**.

Step 4 Confirm the information about the host to be isolated and click **OK**.

After the system displays a message indicating that the operation is successful, click **Finish**. If the host is isolated, its **Operating Status** is displayed as **Isolated**.

Step 5 Cancel the isolation status of the host before using the host if you have rectified the host exception or fault.

1. On MRS Manager, click **Hosts**.
2. Select the check box of the host to be de-isolated and choose **More** > **Cancel Host Isolation**.
3. Click **OK** in the displayed dialog box.

   When **Operation successful** is displayed, click **Finish**. The host is de-isolated successfully, and the value of **Operating Status** becomes **Normal**.
4. Click the name of the de-isolated host to show its status, and click **Start All Roles**.

----End

8.6.4 Modifying the Rack Information of an MRS Cluster Node

All hosts in a large cluster are usually deployed on multiple racks. Hosts on different racks communicate with each other through switches. The network bandwidth between different hosts on the same rack is much greater than that on different racks. In this case, plan the network topology based on the following requirements:

- To improve the communication speed, it is recommended that data be exchanged between hosts on the same rack.
- To improve the fault tolerance capability, distribute processes or data of distributed services on different hosts of multiple racks as dispersedly as possible.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Hadoop uses a file directory structure to represent hosts.

The HDFS cannot automatically determine the network topology of each DataNode in the cluster. You need to set the rack name to identify the rack where the host is located so that the NameNode can draw the network topology of the required DataNodes and back up data of the DataNodes to different racks. Similarly, YARN needs to obtain rack information and allocate tasks to different NodeManagers as required.

If the cluster network topology changes, you need to reallocate racks for hosts on FusionInsight Manager so that related services can be automatically adjusted.

- **Physical rack**: indicates the real rack where the host resides.
- **Logical rack**: indicates the rack name configured on FusionInsight Manager for the host.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Impact on the System

- If the name of the host rack is changed, storage policy for HDFS replicas, YARN task assignment, and storage location of Kafka partitions will be affected. After the modification, you need to restart the HDFS, YARN, and Kafka for the configuration to take effect.
- Improper rack configuration will unbalance loads (including CPU, memory, disk, and network) among nodes in the cluster, which decreases the cluster reliability and stability. Therefore, before allocating racks, take all aspects into consideration and properly set racks.

Rack Allocation Policies

Policy 1: Each logical rack has nearly the same number of hosts.

Policy 2: The name of the logical rack of the host must comply with that of the physical rack to which the host belongs.

Policy 3: If there are only few hosts on a physical rack, combine this physical rack and other physical racks with few hosts into a logical rack, which complies with policy 1. Hosts in two equipment rooms cannot be placed in one logical rack. Otherwise, performance problems may be caused.

Policy 4: If there are lots of hosts on a physical rack, divide these hosts into multiple logical racks, which complies with policy 1. Hosts with great differences should not be placed in the same physical rack. Otherwise, the cluster reliability will be decreased.

Policy 5: You are advised to set **default** or other values for logical racks on the first layer, and the values in the same cluster must be consistent.

Policy 6: The number of hosts in each rack cannot be less than 3.

Policy 7: A cluster can contain at most 50 logical racks. If there are too many logical racks in a cluster, the maintenance is difficult.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Rack Configuration Example

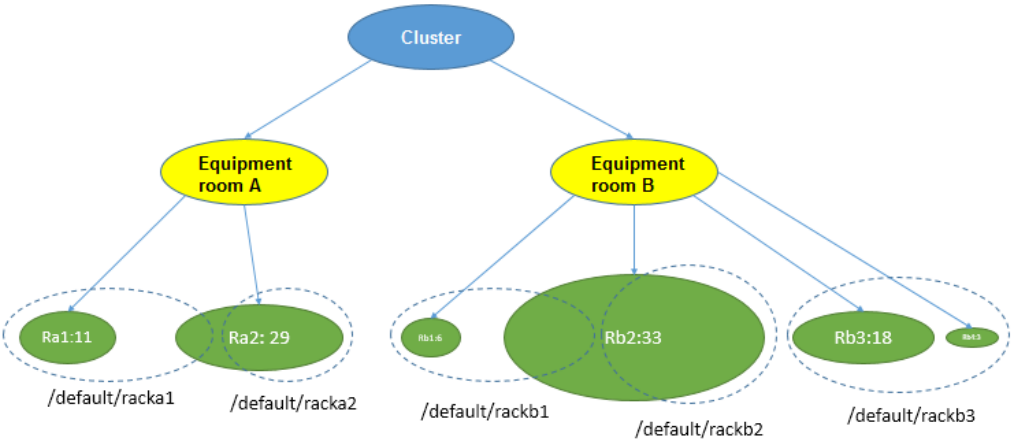
For example, in a cluster, 100 hosts are located in two equipment rooms A and B. A has 40 hosts and B has 60 hosts. In room A, there are 11 hosts on physical rack Ra1 and 29 hosts on physical rack Ra2. In room B, there are six hosts on physical rack Rb1, 33 hosts on physical rack Rb2, 18 hosts on physical rack Rb3, and three hosts on physical rack Rb4.

According to the rack allocation policy, each logical rack contains nearly the same number (for example, 20) of hosts. The allocation details are as follows:

- Logical rack /default/racka1: 11 hosts on physical rack Ra1 and nine hosts on physical rack Ra2
- Logical rack /default/racka2: the remaining 20 hosts (except the nine hosts of logical rack /default/racka1) on physical rack Ra2
- Logical rack /default/rackb1: six hosts on physical rack Rb1 and 13 hosts on physical rack Rb2
- Logical rack /default/rackb2: the remaining 20 hosts on physical rack Rb2
- Logical rack /default/rackb3: 18 hosts on physical rack Rb3 and three hosts on physical rack Rb4

Rack allocation example:

Figure 8-3 Rack division



Configuring the Cluster Node Rack

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts**.

Step 3 Select the check box of the target host.

Step 4 Select **Set Rack** from the **More** drop-down list.

- Set rack names in hierarchy based on the actual network topology. Separate racks from different layers using slashes (/).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Rack naming rules are as follows: */level1/level2/...* The number of levels must be at least 1, and the name cannot be empty. A rack can contain letters, digits, and underscores (_) and cannot exceed 200 characters.

  For example, /default/rack0.
- If the hosts in the rack to be modified contain DataNode instances, ensure that the rack name levels of the hosts where all DataNode instances reside are the same. Otherwise, the configuration fails to be delivered.

Step 5 Click **OK**.

Step 6 Restart the HDFS instances, and YARN and Kafka components for the configuration to take effect.

1. Log in to MRS Manager, choose **Cluster** > **Services** > **HDFS** > **Instances**, select all NameNode instances, choose **More** > **Restart Instance**, verify the password, and click **OK** (No other options need to be selected). Wait until the restart is complete.
2. Restart the DataNode instances by referring to **Step 6.1**.
3. Choose **Cluster** > **Services**, click ··· in the upper right corner of the YARN and Kafka components, respectively, select **Service Rolling Restart**, verify the password, and click **OK** (no other options need to be selected). Wait until the restart is complete.

----End

8.6.5 Managing MRS Cluster Node Groups

Node groups in an MRS cluster are used to manage cluster nodes. A node group is typically composed of ECS nodes that share the same specifications.

You can perform operations on a node group to manage all nodes within the group.

Notes and Constraints

- Master nodes in an MRS cluster cannot be managed by node group.
- A node group can be deleted only when the node type is Task or Core and there are no nodes in the group.
- You can manually add a node group only for a custom cluster of MRS 3.*x*.

Adding a Node Group

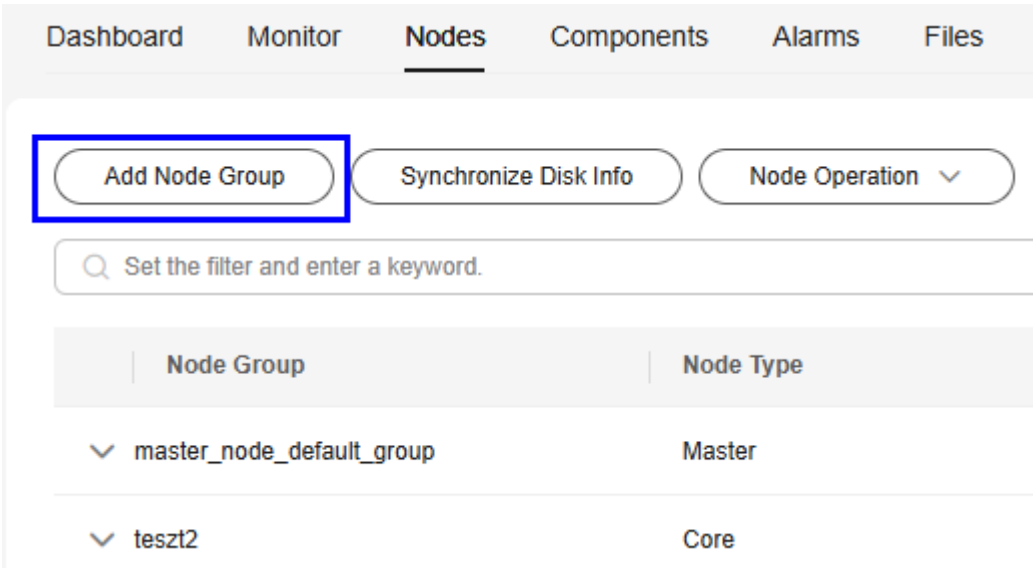
Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click the **Nodes** tab and click **Add Node Group**. The **Add Node Group** page is displayed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 8-4 Add Node Group



Step 4 Set the parameters as required.

Table 8-30 Parameters for adding a node group

Parameter	Description
Name	The node group name must be unique.
CPU Architecture	CPU architecture of an MRS cluster node. The value can be **x86** or **Kunpeng**.
Node Type	• **Task**: A Task node processes data but does not store cluster data. Task nodes support elastic scaling based on service workloads. Data roles that do not store data, such as NodeManager, are deployed on Task nodes. Auto scaling rules can be configured a Task node group. • **Core**: A Core node is an instance that processes and stores data. Component's data storage roles, such as DataNode and Broker, are deployed on Core nodes.
Instance Specifications	Select the instance specifications of the nodes in the current node group. Higher specifications provide stronger data processing and analysis capabilities, but also increase the overall cluster costs. For details about the MRS cluster node specifications, see **MRS Cluster Node Specifications**.
Nodes	Number of nodes in the current node group.
System Disk	Storage type and space of the system disk on a node. You can adjust them as required. For details about the MRS cluster storage, see **Cluster Node Disk Types**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Data Disk	Storage type and space of data disks on a node. A maximum of 10 disks can be added to each Core or Task node. For details about MRS cluster storage, see **Cluster Node Disk Types**.
Disks	Number of data disks on each node in the node group.
Roles	Component roles to be deployed on nodes in the current node group. Only role instances of components that have been installed in the current cluster can be added.
Billing Mode	Billing mode of the current cluster.
Pricing	Price of the node group to be added.

Step 5 Click **OK**.

After the node group is added, you can view it on the node management page.

----End

Deleting a Node Group

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the **Nodes** tab page, locate the row that contains the target node group, and choose **More** > **Delete**.

A node group can be deleted only when the node type is Task or Core and there are no nodes in the group.

Step 4 In the displayed dialog box, click **OK**. Wait until the node group is successfully deleted.

----End

Helpful Links

- For details about how to increase the number of nodes in a node group, see **Scaling Out an MRS Cluster**.
- For details about how to reduce the number of nodes in a node group, see **Scaling In an MRS Cluster**.
- For details about how to configure auto scaling rules for a Task node group, see **MRS Task Node Auto Scaling**.

8.6.6 Adding a Tag to an MRS Cluster/Node

Tags are used to identify clusters/nodes. Adding tags to clusters/nodes can help you identify and manage your resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Cluster tags**: You can add up to 20 tags to a cluster during creation or on the details page of an existing cluster. Updating a cluster tag will synchronize the tag to all nodes in the cluster.
- **Node tags**: You can either use the default tag or add tags to nodes in an MRS cluster when configuring an auto scaling policy. Node tags consume tag quotas. You can view the tags of a node in the **Nodes** tab on the MRS console.
- **Default tags**: An MRS cluster consists of multiple nodes, and each node is an ECS instance that includes EVS disks. After the default tag is enabled, the system automatically generates a default cluster tag and default node tags. These tags are synchronized to the corresponding ECS or EVS disks.

To view node tags, access the **Nodes** tab on the MRS console, and hover over the tag icon of a node in the node list.

- Tag updates in MRS are automatically synchronized to ECSs or EVS disks within the cluster. Changes made directly on the ECS or EVS console will not be reflected in MRS. To ensure tag consistency, do not modify MRS cluster tags on the ECS or EVS console.
- A cluster supports up to 20 tags. When a node in the cluster reaches the maximum number of tags, no additional tags can be added to the cluster.
- When the default tag is enabled, a default tag is added to the cluster and each node, consuming two quotas. By default, a maximum of 20 tags can be added. In this case, up to 18 additional tags can still be added.

If your organization has configured tag strategies for MRS, add tags to clusters/nodes based on the strategies. If a tag does not comply with the tag strategies, the cluster/node may be failed to be created. Contact the organization administrator to learn more about the tag strategies.

A tag consists of a tag key and a tag value. **Table 8-31** provides tag key and value requirements.

Table 8-31 Tag key and value requirements

Parameter	Requirement	Example
Key	A tag key cannot be left blank. A tag key must be unique in a cluster. A tag key can contain a maximum of 128 characters. A tag value cannot contain special characters (=*<>\,|/) or start or end with spaces.	Organization

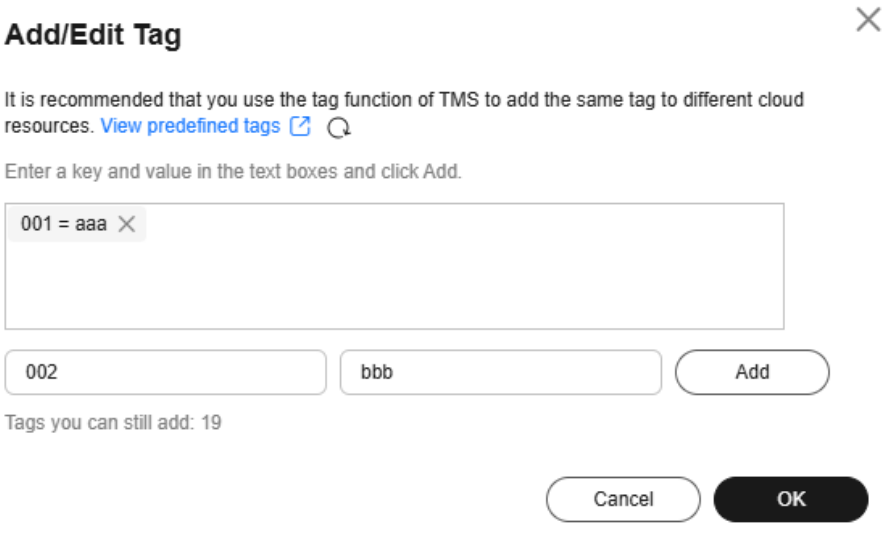
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Requirement	Example
Value	A tag value can contain a maximum of 255 characters. A tag value cannot contain special characters (=*<>\,|/) or start or end with spaces. This parameter can be left blank.	Apache

Adding Tags to a Cluster

- Adding cluster tags during cluster creation
  a. Log in to the MRS console.
  b. Click **Create Cluster**. The page for creating a cluster is displayed.
  c. Click the **Custom Config** tab.
  d. Configure the cluster software and hardware by referring to **Creating a Custom MRS Cluster**.
  e. Enable **Advanced Configuration** and enter the key and value of the tag to be added.
- Adding tags to an existing cluster
  a. Log in to the MRS console.
  b. On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.
  c. Click the **Tags** tab.
  d. Click **Add/Edit Tag**. If this is your first time adding a tag, click **Add Tag**. In the displayed dialog box, enter the key and value of a tag, and click **Add**.

  Figure 8-5 Adding/Editing a Tag



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can also add a tag to a cluster by enabling the default tag. This adds a default tag to the cluster and each node, consuming two tag quotas.

e. Click **OK**.

Adding Tags to a Node

- Adding a default tag to a cluster automatically applies the tag to each node within the cluster. For details, see **Adding tags to an existing cluster**.
- Adding node tags for auto scaling

  If you add a tag when configuring an auto scaling policy, MRS automatically adds the tag to the new nodes and synchronizes the tag to the ECSs and EVS disks.

  a. Log in to the MRS console.
  b. On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.
  c. On the page that is displayed, click the **Auto Scaling** tab.
  d. Click **Edit** on the right of an existing auto scaling policy. In the displayed dialog box, enter the key and value of the tag you want to add, and click **Add**.
  e. Click **OK**.

  When adding node tags through auto scaling:

  – You need to enable the auto scaling policy and configure scale-out rules. Otherwise, the node tags will not take effect.
  – If tag quotas are insufficient, delete the cluster tag or modify an existing tag in the auto scaling policy, and then enable the policy.
  – You cannot add tags to auto scaling policies of resource pools.

Searching for Target Clusters by Tags

On the **Active Clusters** page, search for the target cluster by tag key or tag value.

1. Log in to the MRS console.
2. In the upper right corner of the **Active Clusters** page, click **Search by Tag** to access the search page.
3. Enter the tag of the cluster to be searched.

   You can select a tag key or tag value from their drop-down lists. When the tag key or tag value is exactly matched, the system can automatically locate the target cluster. If you enter multiple tags, their intersections are used to search for the cluster.
4. Click **Search**.

   The system searches for the target cluster by tag key or value.

Managing Tags

You can view, add, and delete tags on the **Tags** tab page of the cluster.

1. Log in to the MRS console.
2. On the **Active Clusters** page, click the name of a cluster for which you want to manage tags.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The cluster details page is displayed.

3. Click the **Tags** tab and view, add, and delete tags on the tab page.
   - View

     On the **Tags** tab page, you can view details about tags of the cluster, including the number of tags and the key and value of each tag.
   - Add

     Click **Add/Edit Tag**. If this is your first time adding a tag, click **Add Tag**. In the displayed dialog box, enter the key and value of a tag, and click **OK**.
   - Delete

     Locate the row that contains the tag you want to delete and click **Delete** in the **Operation** column. In the displayed **Delete Tag** dialog box, and click **Yes**.

8.6.7 Configuring Bootstrap Actions for an MRS Cluster Node

8.6.7.1 MRS Bootstrap Action Overview

You can run bootstrap actions to install additional third-party software, modify the cluster running environment, and perform other customizations.

Bootstrap actions can execute scripts on specified nodes before or after the first startup of cluster components. You can only manually run the third-party component installation script on the node to install a running cluster component.

If you choose to run bootstrap actions when scaling out a cluster, the bootstrap actions will be run on the newly added nodes in the same way.

If you use auto scaling, MRS allows you to customize automation scripts for auto scaling. Automation scripts can be executed before and after auto scaling and automatically adapt to service workload changes, freeing you from manual operations. In addition, automation scripts, which are fully customizable and can be scheduled for execution at specific times, can significantly enhance auto scaling flexibility and cater to individual needs.

WARNING

- MRS 2.*x* and earlier versions: Bootstrap action scripts must be executed as user **root**. Improper use of the scripts may affect cluster availability. Exercise caution when using the scripts. You can run the **su -** *xxx* command to switch users in the script.
- MRS 3.*x* or later: Bootstrap action scripts must be executed as user **omm**. Improper use of the scripts may affect cluster availability. Exercise caution when using the scripts. You can run the **su -** *xxx* command to switch users in the script.

MRS determines the result based on the return code after the execution of the bootstrap action script. If the return code is **0**, the script is executed. If the return code is not **0**, the execution fails. If a bootstrap action script fails to be executed

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

on a node, the corresponding boot script will fail to be executed. In this case, you can set **Action upon Failure** to choose whether to continue to execute the subsequent scripts.

- Example 1: If you set **Action upon Failure** to **Continue** for all scripts during cluster creation, all the scripts will be executed regardless of whether they are successfully executed, and the startup process will be complete.
- Example 2: If a script fails to be executed and **Action upon Failure** is set to **Stop**, subsequent scripts will not be executed and cluster creation or scale-out will fail.

You can add a maximum of 18 bootstrap actions, which will be executed before or after the cluster component is started in the order you specified. The bootstrap actions performed before or after the component startup must be completed within 60 minutes. Otherwise, the cluster creation or scale-out will fail.

8.6.7.2 Preparing the Bootstrap Action Script for an MRS Node

Currently, bootstrap actions support Linux shell scripts only. Script files must end with **.sh**.

Follow these steps to prepare the bootstrap action script for the MRS node:

Step 1 Upload the required installation packages to the OBS file system.

Before compiling a script, you need to upload all required installation packages, configuration packages, and relevant files to the OBS file system in the same region.

Different regions are isolated from each other. Therefore, MRS VMs cannot download OBS files from other regions.

Step 2 Use a script for downloading files from the OBS file system.

You can specify the file to be downloaded from OBS in the script. To upload a file to a private file system, you need to use **hadoop fs** to download the file.

To download the **myfile.tar.gz** file from the **obs://yourbucket/** directory to your local host and decompress it to the **/your-dir** directory, run the following commands:

```
source /opt/Bigdata/client/bigdata_env;hadoop fs -D fs.obs.endpoint=<obs-endpoint> -D fs.obs.access.key=<your-ak> -D fs.obs.secret.key=<your-sk> -copyToLocal obs://yourbucket/myfile.tar.gz ./
mkdir -p /<your-dir>
tar -zxvf myfile.tar.gz -C /<your-dir>
```

- **/opt/Bigdata/client** indicates the client path. Replace it with the actual one.
- The Hadoop client has been preinstalled on the MRS node. You can run the **hadoop fs** command to download data from or upload data to OBS.
- Obtain the OBS endpoint of each region. For details, see **Regions and Endpoints**.
- Commands carrying authentication passwords pose security risks. Disable historical command recording before running such commands to prevent information leakage.

Step 3 Upload the script to the OBS file system.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After script compilation, upload the script to the OBS file system in the same region. At the time you specify, each node in the cluster downloads the script from OBS and executes it as user **root**.

----End

8.6.7.3 Adding MRS Node Bootstrap Actions and Installing Third-Party Software

Prerequisites

You have prepared the bootstrap action script by referring to **Preparing the Bootstrap Action Script for an MRS Node**.

Adding Bootstrap Actions

Bootstrap actions can be added during or after cluster creation. You can also configure the automatic execution script on the auto scaling page.

Adding Bootstrap Actions During Cluster Creation

Step 1 Log in to the MRS console.

Step 2 Click **Create Cluster**. The page for creating a cluster is displayed.

Step 3 Click **Custom Config**.

Step 4 Configure the cluster software and hardware by referring to **Creating a Custom MRS Cluster**.

Step 5 Enable **Advanced Configuration**, click **Add** in the **Bootstrap Action** area, and configure the following parameters.

Table 8-32 Parameter description

Parameter	Description	Example Value
Name	Name of the bootstrap action script. The name must be unique within a cluster, but can be the same across different clusters. The value can contain only digits, letters, spaces, hyphens (-), and underscores (_) and must not start with a space. The value can contain 1 to 64 characters.	bootstrap_0

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Script Path	Path of the script. The value can be an OBS file system path or a local VM path. **The path must be unique within a cluster, but can be the same across different clusters.** ● An OBS file system path must start with **obs://** and end with **.sh**, for example, **obs://mrs-samples/*xxx*.sh**. ● A local VM path must start with a slash (/) and end with **.sh**.	obs://mrs-samples/xxx.sh
Parameter	Bootstrap action script parameters	-
Execution Node	Select a type of the node where the bootstrap action script is executed.	-
Executed	Select the time when the bootstrap action script is executed. **To add bootstrap actions for a running cluster, you can only manually run the third-party component installation script on the node.** ● Before initial component start ● After initial component start	After initial component start
Action upon Failure	Whether to continue to execute subsequent scripts and create a cluster after the script fails to be executed. You are advised to set this parameter to **Continue** in the debugging phase so that the cluster can continue to be installed and started no matter whether the bootstrap action is successful.	Continue
Run as root	Whether to escalate the permission to user **root**. (This parameter is available only in clusters of MRS 3.1.5, MRS 3.3.0, and later versions.) If the bootstrap action requires root user operations, enable this function, or the bootstrap action may fail to execute.	-

Step 6 Click **OK**.

After the bootstrap action is added, you can edit, clone, or delete it in the **Operation** column.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Adding Bootstrap Actions for an Existing Cluster

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the page that is displayed, click the **Bootstrap Actions** tab.

Step 4 Click **Add** and set parameters as prompted.

Table 8-33 Parameter description

Parameter	Description	Example Value
Name	Name of a bootstrap action script The name must be unique within a cluster. You can set the same name for different clusters. The value can contain only digits, letters, spaces, hyphens (-), and underscores (_) and must not start with a space. The value can contain 1 to 64 characters.	bootstrap_0
Script Path	Script path. The value can be an OBS file system path or a local VM path. **The path must be unique within a cluster, but can be the same across different clusters.** • An OBS file system path must start with **obs://** and end with **.sh**, for example, **obs://mrs-samples/*xxx*.sh**. • A local VM path must start with a slash (/) and end with **.sh**.	obs://mrs-samples/xxx.sh
Execution Node	Select a type of the node where the bootstrap action script is executed.	-
Active Master	Whether the task will be executed only on the active Master node. This parameter is valid only when the Master node is selected as the execution node. By default, the function is disabled, indicating that the command will be executed on all Master nodes.	-
Parameter	(Optional) Bootstrap action script parameter.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Executed	Select the time when the bootstrap action script is executed. **To add bootstrap actions for a running cluster, you can only manually run the third-party component installation script on the node.** ● Before initial component start ● After initial component start	After initial component start
Action upon Failure	Whether to continue to execute subsequent scripts and create a cluster after the script fails to be executed. You are advised to set this parameter to **Continue** in the debugging phase so that the cluster can continue to be installed and started no matter whether the bootstrap action is successful.	Continue
Run as root	Whether to escalate the permission to user **root**. (This parameter is available only in clusters of MRS 3.1.5, MRS 3.3.0, and later versions.) If the bootstrap action requires root user operations, enable this function, or the bootstrap action may fail to execute.	-

Step 5 Click **OK** to save the configuration.

Step 6 Click **Yes**. After a bootstrap action is added, you can modify or delete it in the bootstrap action list.

----End

Adding an Automatic Script for Auto Scaling Nodes

If you use auto scaling, MRS allows you to customize automation scripts for auto scaling. Automation scripts can be executed before and after auto scaling and automatically adapt to service load changes, all of which eliminates manual operations.

- You can set an automation script so that it can automatically run on cluster nodes when auto scaling is triggered.
- Before using automation scripts, upload them to a cluster VM or OBS file system in the same region as the cluster. The automation scripts uploaded to the cluster VM can be executed only on the existing nodes. If you want to make the automation scripts run on the new nodes, upload them to the OBS file system.

Step 1 Log in to the MRS console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click **Auto Scaling**.

Step 4 Click **Configure Automation Script**.

Step 5 Click **Create**. Set parameters by referring to **Table 8-33**.

For the parameters to be configured, **Executed** can be set to **Before scale-out**, **After scale-out**, **Before scale-in**, or **After scale-in**.

Step 6 Click **OK** to save the automation script configurations.

----End

Helpful Links

- You can configure auto scaling for MRS cluster nodes to enable the cluster to dynamically adjust the number of nodes in it based on data volume changes. For details, see **MRS Task Node Auto Scaling**.
- If bootstrap actions are added during cluster creation and you have configured **Run Bootstrap Action** for cluster node scaling settings, the nodes added by auto scaling will execute all the bootstrap action scripts added during cluster creation. For details, see **Scaling Out an MRS Cluster**.

8.6.7.4 Viewing the Bootstrap Action Execution Records of an MRS Node

You can view the execution result of the bootstrap operation on the **Bootstrap Action** page.

Viewing the Execution Result

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Bootstrap Action**. Information about the bootstrap actions added during cluster creation is displayed.

- You select **Before initial component start** or **After initial component start** in the upper right corner to query information about the related bootstrap actions.
- The last execution result is listed here. For a newly created cluster, the records of bootstrap actions executed during cluster creation are listed. If a cluster is expanded, the records of bootstrap actions executed on the newly added nodes are listed.

----End

Viewing Execution Logs

To view the run logs of a bootstrap action, set **Action upon Failure** to **Continue** when adding the bootstrap action, and then log in to each node to view the run logs. For clusters of MRS 3.x or later, run logs are stored in the **/var/log/Bigdata/**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Bootstrap directory. For clusters of MRS 2.x, run logs are stored in the **/var/log/Bootstrap** directory.

If you add bootstrap actions before and after component start, you can distinguish bootstrap action logs of the two phases based on the timestamps.

You are advised to print logs in detail in the script so that you can view the detailed run result. MRS redirects the standard output and error output of the script to the log directory of the bootstrap action.

8.7 Managing the MRS Cluster Client

8.7.1 Updating the MRS Cluster Client After the Server Configuration Expires

An MRS cluster provides a client for you to connect to a server, view task results, or manage data. If you modify service configuration parameters on Manager and restart the service, you need to download and install the client again or use the configuration file to update the client.

For clusters of MRS 2.*x* or earlier, the original client is stored in the **/opt/client** directory of all nodes in the cluster by default when you create a cluster. After the cluster is created, only the client of a Master node can be directly used. To use the client of a Core node, you need to update the client configuration file first.

Updating Client Configurations (Version 3.x or Later)

Method 1

Step 1 Log in to MRS Manager by referring to **Accessing MRS Manager**.

- Versions earlier than MRS 3.3.0: Choose **Cluster** > **Dashboard** > **More** > **Download Client**.
- MRS 3.3.0 or later: On the Manager homepage, click **Download Client**.

Step 2 Choose **More** > **Download Client** > **Configuration Files Only**.

The generated compressed file contains the configuration files of all services.

Step 3 Set the parameters.

For versions earlier than MRS 3.5.0:

Determine whether to generate a configuration file on the cluster node.

- If yes, select **Save to Path**, and click **OK** to generate the client file. By default, the client file is generated in **/tmp/FusionInsight-Client** on the active management node. You can also store the client file in other directories, and user **omm** has the read, write, and execute permissions on the directories. Then go to **Step 4**.
- If no, click **OK**, specify a local save path, and download the complete client. Wait until the download is complete and go to **Step 4**.

For MRS 3.5.0 and later versions:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Set the **Select Download Location** parameter as needed.
   - **Server**: Download the file to the active OMS node of the cluster.
   - **Browser**: Download the file to the local computer.
   - **Remote Node:** Download the file to a node other than the active OMS node.
2. Set **Save to Path** as prompted. If you select **Browser** as the download location, left this parameter blank.
3. If you select **Remote Node**, configure the required parameters by referring to **Table 8-34**.

Table 8-34 Parameters for downloading to a remote node

Parameter	Description	Example Value
Host IP Address	IP address of the remote node. The platform type of the remote node must be the same as that of the downloaded client. Otherwise, the client may fail to be installed.	x.x.x.x
Host Port	Host port of the remote node.	22
Username	Username for logging in to the remote node. The user must have the write permission for the path.	xxx
Authentication Method	You can choose one of the following methods: - **Password**: Use the password for login. - **SSH private keys**: Use SSH private keys for login. - **None**: To use this method, passwordless login needs to be enabled.	Password
Password	Enter the password when **Authentication Method** is set to **Password**. Password for logging in to the remote node.	xxx
SSH private keys	This parameter is mandatory when **Authentication Method** is set to **SSH private keys**. Click **Add** and select the SSH private key file to upload.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Auto Deployment	Whether to enable auto deployment. This parameter is mandatory when **Select Client Type** is set to **Complete Client**. – If you set this parameter to **yes**, the client is automatically installed and deployed on the current node. – If you set this parameter to **no**, the client will not be automatically installed and deployed. You need to manually install the client after it is downloaded.	Yes
Deployment Path	This parameter is mandatory when **Auto Deployment** is set to **Yes**. If only the configuration file is downloaded, this parameter will not be displayed. The deployment path must be empty if it already exists on the remote node. Otherwise, it will be created automatically. The path also requires operate and write permissions.	/opt/testclient

Step 4 Use WinSCP to save the compressed file to the client installation directory, for example, **/opt/hadoopclient**, as the client installation user.

Step 5 Decompress the software package.

Run the following commands to go to the directory where the client is installed, and decompress the file to a local directory. For example, the downloaded client file is **FusionInsight_Cluster_1_Services_Client.tar**.

```
cd /opt/hadoopclient
tar -xvf FusionInsight_Cluster_1_Services_Client.tar
```

Step 6 Verify the software package.

Run the following command to verify the decompressed file and check whether the command output is consistent with the information in the **sha256** file.

```
sha256sum -c FusionInsight_Cluster_1_Services_ClientConfig_ConfigFiles.tar.sha256
```

The following information is displayed:

```
FusionInsight_Cluster_1_Services_ClientConfig_ConfigFiles.tar: OK
```

Step 7 Decompress the package to obtain the configuration file.

```
tar -xvf FusionInsight_Cluster_1_Services_ClientConfig_ConfigFiles.tar
```

Step 8 Run the following command in the client installation directory to update the client using the configuration file:

```
sh refreshConfig.sh Client installation directory Directory where the configuration file is stored
```

For example, run the following command:

```
sh refreshConfig.sh /opt/hadoopclient /opt/hadoopclient/
FusionInsight_Cluster_1_Services_ClientConfig_ConfigFiles
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the following information is displayed, the configurations have been updated successfully.

```
Succeed to refresh components client config.
```

----End

Method 2:

Step 1 Log in to the client installation node as user **root**.

Step 2 Go to the client installation directory, for example, **/opt/hadoopclient** and run the following commands to update the configuration file:

```
cd /opt/hadoopclient
sh autoRefreshConfig.sh
```

Step 3 Enter the username and password of the FusionInsight Manager administrator and the floating IP address of OMS.

To obtain the floating IP address of OMS, log in to the Master2 node remotely, and run the **ifconfig** command. In the command output, **eth0:wsom** indicates the floating IP address of OMS. Record the value of **inet**. If the floating IP address of OMS cannot be queried on the Master2 node, switch to the Master1 node to query and record the floating IP address. If there is only one Master node, query and record the IP address on the Master node.

Step 4 Enter the names of the components whose configuration needs to be updated. Use commas (,) to separate the component names. Press **Enter** to update the configurations of all components if necessary.

If the following information is displayed, the configurations have been updated successfully.

```
Succeed to refresh components client config.
```

----End

Method 3 (for MRS 3.5.0 and later):

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Client Management**. On the displayed page, check whether there is the client you want to operate.

- If yes, go to **Step 3**.
- If no, click **Add**, configure information about the client, and click **OK**. Go to **Step 3**.

Step 3 Click **Update** in the **Operation** column of the client you want to update.

If you need to update clients in batches, select the target clients (use the same installation user) and choose **More** > **Batch Update**.

Step 4 Select the checkbox for risk confirmation, configure the information in **Table 8-35**, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

WARNING

During a configuration update, clients may not function properly and client services may be interrupted.

Table 8-35 Parameters for updating the client configuration

Parameter	Example Value	Description
Authentication Method	Password	You can choose one of the following methods: • **Password**: Use the password for login. • **SSH private keys**: Use SSH private keys for login. • **None**: To use this method, passwordless login needs to be enabled.
Password/SSH Private Keys	xxx	• Enter the password if **Authentication Method** is set to **Password**. • Upload the private key file if **Authentication Method** is set to **SSH private keys**. Click **Select File** and select a local file to upload.
Host Port	22	Host port of the node where the client is to be updated
Installation Package Path	/tmp/FusionInsight-Client/	Path for storing the installation package of the updated client If there already is a client file in the path, it will be overwritten. For a remote node, write permission for the path is required.

----End

Updating Client Configurations (Version 2.x or Earlier)

Method 1

Step 1 Log in to MRS Manager and choose **Services**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Click **Download Client**.

Set **Client Type** to **Only configuration files**, **Download To** to **Server**, and click **OK** to generate the client configuration file. The generated file is saved in the **/tmp/MRS-client** directory on the active management node by default. You can customize the file path.

Step 3 Query and log in to the active Master node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 If you use the client in the cluster, run the following command to switch to user **omm**. If you use the client outside the cluster, switch to user **root**.

```
sudo su - omm
```

Step 5 Go to the client directory.

```
cd Client installation directory
```

Step 6 Update client configurations.

```
sh refreshConfig.sh Client installation directoryFull path of the client configuration file package
```

Example:

```
sh refreshConfig.sh /opt/Bigdata/client /tmp/MRS-client/MRS_Services_Client.tar
```

If the following information is displayed, the configurations have been updated successfully.

```
ReFresh components client config is complete.
Succeed to refresh components client config.
```

----End

Method 2:

Step 1 After the cluster is installed, run the following command to switch to user **omm**. If you use the client outside the cluster, switch to user **root**.

```
sudo su - omm
```

Step 2 Go to the client directory.

```
cd {Client installation directory}
```

Step 3 Run the following command and enter the name of an MRS Manager user with the download permission and its password (for example, the username is **admin** and the password is the one set during cluster creation) as prompted to update client configurations.

```
sh autoRefreshConfig.sh
```

Step 4 After the command is executed, the following information is displayed, where *XXX* indicates the name of the component installed in the cluster. To update client configurations of all components, press **Enter**. To update client configurations of some components, enter the component names and separate them with commas (,).

```
Components "xxx" have been installed in the cluster. Please input the comma-separated names of the components for which you want to update client configurations. If you press Enter without inputting any component name, the client configurations of all components will be updated:
```

If the following information is displayed, the configurations have been updated:

```
Succeed to refresh components client config.
```

If the following information is displayed, the username or password is incorrect.

```
login manager failed,Incorrect username or password.
```

- This script automatically connects to the cluster and invokes the **refreshConfig.sh** script to download and update the client configuration file.
- By default, the client uses the floating IP address specified by **wsom=xxx** in the **Version** file located in the installation directory to update the client configurations. To update the configuration file of another cluster, change the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

value of **wsom=xxx** in the **Version** file to the floating IP address of the target cluster before performing this step.

----End

8.7.2 Viewing the Installed MRS Cluster Client

FusionInsight Manager supports unified management of cluster client installation information. After a user downloads and installs a client, FusionInsight Manager automatically records information about the installed (registered) client to facilitate query and management.

In addition, you can manually add or modify the information about clients that are not automatically registered, for example, clients installed in earlier versions.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Viewing the Installed Cluster Client

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** and choose **Client Management** to view information about clients installed in the cluster.

You can view the IP address, installation path, component list, registration time, and installation user of the node where the client is located.

When the client is downloaded and installed in the cluster of the latest version, the client information is automatically registered.

Step 3 To manually add information about an installed client, click **Add** and manually add the IP address, installation path, user, platform information, and registration information of the client as prompted.

Configure the client information and click **OK**.

You can also manually modify or delete the information about the manually registered client.

Step 4 On the **Client Management** page, click **Export All** to export information about all registered clients to the local PC.

On the **Client Management** page, only components with clients are displayed in the component list. Therefore, components without clients or with special clients are not displayed.

----End

8.7.3 Batch Upgrading MRS Cluster Clients

The client package downloaded from FusionInsight Manager contains the client batch upgrade tool. When multiple clients need to be upgraded after the cluster

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

upgrade or scale-out, you can use this tool to upgrade the clients in batches with a few clicks. In addition to upgrading clients in batches, this tool also provides the lightweight function of updating the **/etc/hosts** files on the nodes where the clients are located in batches.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Preparing for the Batch Upgrade

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Download the client on FusionInsight Manager.

- Versions earlier than MRS 3.3.0: Choose **Cluster** > **Dashboard** > **More** > **Download Client**.
- MRS 3.3.0 or later: On the Manager homepage, click **Download Client**.

Download the full client to the specified directory on the server. For details, see **Installing an MRS Cluster Client**.

Decompress the downloaded client package and find the **batch_upgrade** directory, for example, **/tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_ClientConfig/batch_upgrade**.

Step 3 Choose **Cluster** > **Client Management**. On the **Client Management** page, click **Export All** to export all client information to the local PC.

Step 4 Decompress the exported client information and upload the **client-info.cfg** file to the **batch_upgrade** directory.

Step 5 Perform the following operations to supplement the missing passwords in the **client-info.cfg** file:

Run the following command to add a user password:

```
vi client-info.cfg
```

Example:

```
clientIp,clientPath,user,password
10.10.10.100,/home/omm/client /home/omm/client2,omm,Password
```

The fields in the configuration file are as follows:

- **clientIp**: indicates the IP address of the node where the client is located.
- **clientPath**: indicates the client installation path. Multiple paths are separated by spaces. Note that the path cannot end with a slash (/).
- **user**: indicates the username of the node.
- **password**: indicates the user password of the node.

If the execution fails, view the **node.log** file in the **work_space/log_*XXX*** directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

WARNING

Configuration files containing authentication passwords pose security risks. Delete such files after configuration or store them securely.

----End

Procedure

Step 1 Log in to the client download node as the user who wants to install the client.

Step 2 Perform the upgrade.

```
sh client_batch_upgrade.sh -u -f /tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_Client.tar -g /tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_ClientConfig/batch_upgrade/client-info.cfg
```

WARNING

You are advised to delete the **client-info.cfg** file as soon as possible after the upgrade because the password has been configured.

Step 3 After the upgrade is complete, verify the upgrade result by running the following command:

```
sh client_batch_upgrade.sh -c
```

Step 4 If the client does not function properly after the upgrade, run the following command to roll back the client:

```
sh client_batch_upgrade.sh -s
```

- The client batch upgrade tool moves the original client to the backup directory, and then installs the client using the package specified by the **-f** parameter. Therefore, if the original client contains customized content, manually save the customized content from the backup directory or move it to the client directory after the upgrade, before running the **-c** command.

  The backup path on the client is *{Original client path}*-**backup**.
- The **-u** command is the prerequisite for the **-c** and **-s** commands. You can run the **-c** command to commit the upgrade or the **-s** command to perform a rollback only after the **-u** command is executed to perform an upgrade.
- You can execute the **-u** command multiple times. Each execution upgrades only the clients that failed previously, while successfully upgraded clients are skipped.
- The client batch upgrade tool also supports the clients of earlier versions.
- When upgrading a client installed by a non-root user, ensure that the user has the read and write permissions on both the client directory and its parent directory on the target node. Otherwise, the upgrade will fail.
- The client package specified by the **-f** parameter must be a full client package. Client packages containing only a single component or a subset of components cannot be used as input.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Updating the hosts Files in Batches

Step 1 Complete all preparations by referring to **Preparing for the Batch Upgrade**.

Step 2 Check whether the user configured for the node where the **/etc/hosts** files need to be updated is **root**.

- If yes, go to **Step 3**.
- If no, change the user to **root** and go to **Step 3**.

Step 3 Update the **/etc/hosts** files on the nodes where the client is located in batches.

```
sh client_batch_upgrade.sh -r -f /tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_Client.tar -
g /tmp/FusionInsight-Client/FusionInsight_Cluster_1_Services_ClientConfig/batch_upgrade/client-info.cfg
```

- When you batch update the **/etc/hosts** files, the specified client package can be either a complete client package or a package containing only configuration files (recommended).
- The user configured for the host where the **/etc/hosts** files need to be updated must be **root**. Otherwise, the update fails.

----End

8.8 Managing MRS Cluster Jobs

8.8.1 Stopping and Deleting an MRS Cluster Job

You can manually stop a running MRS job on the console.

After a job is executed, you can delete it if you do not need to view the job information.

Stopping an MRS Cluster Job

Spark SQL jobs cannot be stopped. After a job is stopped, its status changes to **Terminated** and the job cannot be executed again.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click **Jobs**.

Step 4 Select a running job, and choose **More** > **Stop** in the **Operation** column.

The job status changes from **Running** to **Terminated**.

----End

Deleting an MRS Cluster Job

You can delete a single job or multiple jobs in batches. A deleted job cannot be restored. Therefore, exercise caution when deleting a job.

Step 1 Log in to the MRS console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click **Jobs**.

Step 4 Choose **More** > **Delete** from the **Operation** in the row of the target job to be deleted.

This step deletes only a single job.

Step 5 Select multiple jobs and click **Delete** on the upper left of the job list.

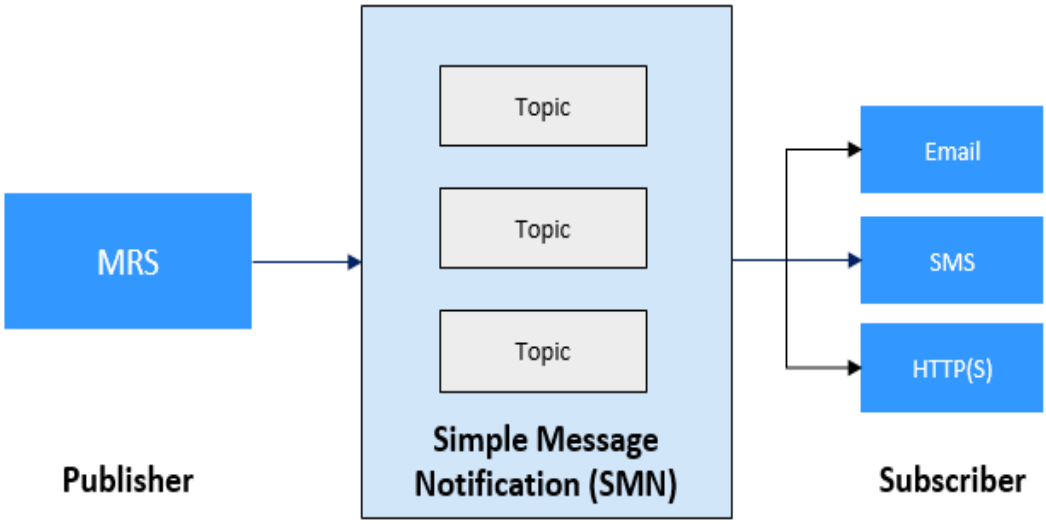
You can delete one, multiple, or all jobs.

----End

8.8.2 Configuring Notification Rules for MRS Jobs

Simple Message Notification (SMN) is a reliable and flexible large-scale message notification service. It enables you to efficiently send messages to various endpoints, such as email addresses and phone numbers. MRS uses SMN to offer a publish/subscribe model to achieve one-to-multiple message subscriptions and notifications through different subscription endpoints (SMSs and emails).

Figure 8-6 MRS cluster message notification



You can configure job notification rules to receive notifications immediately upon a job execution success or failure.

Step 1: Creating an SMN Topic and Adding a Subscription

A topic is used to publish messages and subscribe to notifications. It serves as a message sending channel, where publishers and subscribers can interact with each other.

Step 1 Log in to the management console.

Step 2 Click **Service List**. Under **Management & Governance**, click **Simple Message Notification**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 In the navigation pane, choose **Topic Management** > **Topics**.

Step 4 Click **Create Topic**.

Table 8-36 Parameters for creating an SMN topic

Parameter	Description	Example Value
Topic Name	Topic name, which: • Contains only letters, digits, hyphens (-), and underscores (_) and must start with a letter or digit. • Contains 1 to 255 characters. • Must be unique and cannot be modified after the topic is created.	mrs-job
Display name	Message sender name, which can contain up to 192 bytes. After you specify a display name, the sender in email messages will be presented as *Display name*<**username@example.com**>. Otherwise, the sender will be **username@example.com**.	mrs-job-failed
Enterprise Project	Name of an enterprise project. An enterprise project allows you to group cloud resources under specific projects, making it easier to manage cloud resources and users. The default name is **default**. This parameter is for enterprise users.	default

Step 5 Click **OK**.

After a topic is created, the system generates a topic URN, which uniquely identifies the topic and cannot be changed.

The created topic is displayed in the topic list. You can click the topic name to view its details.

Step 6 Locate the topic to which you want to add a subscription, click **More** in the **Operation** column, and select **Add Subscription**.

Endpoint indicates the address of the subscription endpoint. You can add multiple SMS and email endpoints at a time, one in each line. You can enter a maximum of 10 endpoints.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-37 Parameters for adding a subscription

Parameter	Description	Example Value
Protocol	Protocol supported by the subscription endpoints. For example, you can receive messages through SMS messages or emails.	Email
Method	The following two options are available: • **Add subscriptions** • **Import subscriptions**: Select subscribers whose subscription has been confirmed and import the subscription data.	Add subscriptions
Endpoint	Subscription endpoint. You can enter up to 10 SMS, email, HTTP, or HTTPS endpoints, one in each line.	username@example.com

Step 7 Click **OK**. The new subscription is displayed in the subscription list.

Once you have added a subscription, SMN will send a confirmation message to the subscription endpoint, which includes a link for confirming the subscription. The subscription confirmation link is valid for 48 hours. Make sure to confirm your subscription on your mobile phone, mailbox, or other endpoints within this time frame.

The subscription endpoint can receive messages only after the subscription status changes to **Confirmed**.

----End

Step 2: Configuring Job Message Notification for an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 Click the name of a cluster to go to the cluster details page.

Step 3 Click the **Alarms** tab, and choose **Notification Rules** > **Add Notification Rule**.

Step 4 Configure a notification rule for sending job execution results to subscribers, and click **OK**.

Table 8-38 Parameters of adding a notification rule

Parameter	Description	Example Value
Rule Name	User-defined notification rule name. Only digits, letters, hyphens (-), and underscores (_) are allowed.	default_alert_rule1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Message Notification	If you enable this function, subscription messages will be sent to subscribers.	Enabled
Topic Name	Select the topic created in **Step 1: Creating an SMN Topic and Adding a Subscription** or click **Create Topic** to create a topic.	mrs-job
Notification Type	Select the type of messages to be reported to SMN. For job information, select **Event**.	Event
Subscription Items	Select events you want to subscribe to. 1. Click next to **Suggestion**. 2. Click next to **Manager**. 3. Select events related to job running, for example, **Job Running Succeeded** and **Job Running Failed**.	Job Running Failed

Step 5 After message notifications are configured for jobs, if a job fails to run, the relevant endpoint can receive the detailed notification.

----End

Helpful Links

- Alarms and events of the MRS cluster can be pushed using SMN. For details, see **Configuring Notifications for MRS Cluster Alarms and Events**.
- For details about how to submit and configure jobs on the MRS console, see **Submitting a Job in an MRS Cluster**.

8.9 Managing MRS Cluster Tenants

8.9.1 Introduction to MRS Multi-Tenancy

Overview

- **Background**:

  Modern enterprises' data clusters are becoming more and more centralized and cloud-based. Enterprise-class big data clusters must meet the following requirements:

  – Carry data of different types and formats and run jobs and applications of different types (such analysis, query, and stream processing).

  – Isolate data of a user from that of another user who has demanding requirements on data security, such as a bank or government institute.

  The preceding requirements bring the following challenges to the big data clusters:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Proper allocation and scheduling of resources to ensure stable operating of applications and jobs.
  - Strict access control to ensure data and service security.

  Multi-tenancy isolates the resources of a big data cluster into resource sets. Users can lease desired resource sets to run applications and jobs and store data. In a big data cluster, multiple resource sets can be deployed to meet diverse requirements of multiple users.

  MRS provides a complete enterprise-class big data multi-tenant solution.

- **Multi-tenancy:**

  An MRS cluster provides various resources and services that can be shared among organizations, departments, or applications. The cluster provides logical entities, that is, tenants, to use these resources and services. Currently, only the analysis cluster supports tenant management.

  A mode involving different tenants is called multi-tenant mode. Multi-tenancy refers to multiple resource sets (a resource set is a tenant) in the MRS cluster and is able to allocate and schedule resources. The resources include computing resources and storage resources. The MRS cluster offers multi-tenancy, supporting a layered tenant model and enabling dynamic tenant creation or deletion to isolate resources. It dynamically allocates and configures compute and storage resources for each tenant.

  The computing resources indicate tenants' Yarn task queue resources. The task queue quota can be modified, and the task queue usage status and statistics can be viewed.

  The storage resources can be stored on HDFS. You can add and delete the HDFS storage directories of tenants, and set the quotas of file quantity and the storage space of the directories.

  Tenants can create and manage tenants in a cluster based on service requirements.

  - Roles, computing resources, and storage resources are automatically created when tenants are created. By default, all permissions of the new computing resources and storage resources are allocated to a tenant's roles.
  - Permissions to view the current tenant's resources, add a subtenant, and manage the subtenant's resources are granted to the tenant's roles by default.
  - After you have modified the tenant's computing or storage resources, permissions of the tenant's roles are automatically updated.

  MRS supports a maximum of 512 tenants. The default tenants created by the system include **default**. Tenants that are in the topmost layer with the default tenant are called level-1 tenants.

- **Resource pool:**

  YARN task queues support only one scheduling policy, namely label based scheduling. This policy allows YARN task queues to select NodeManagers with specific node labels, ensuring that tasks run on designated nodes and utilize target hardware resources. For example, YARN tasks requiring a large memory capacity can run on nodes labeled as large memory resources, preventing poor service performance.

  In an MRS cluster, tenants logically partition nodes to form a resource pool with multiple NodeManagers. YARN task queues can be associated with

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

specified resource pools by configuring queue capacity policies, ensuring efficient and independent resource utilization in the resource pools.

MRS supports a maximum of 50 resource pools. The system has a **default** resource pool.

- **Advantages:**
  - Proper resource configuration and isolation

    The resources of a tenant are isolated from those of another tenant. The resource use of a tenant does not affect other tenants. This mechanism ensures that each tenant can configure resources based on service requirements, improving resource utilization.
  - Resource consumption measurement and statistics

    Tenants are system resource applicants and consumers. System resources are planned and allocated based on tenants. Resource consumption by tenants can be measured and collected.
  - Assured data security and access security

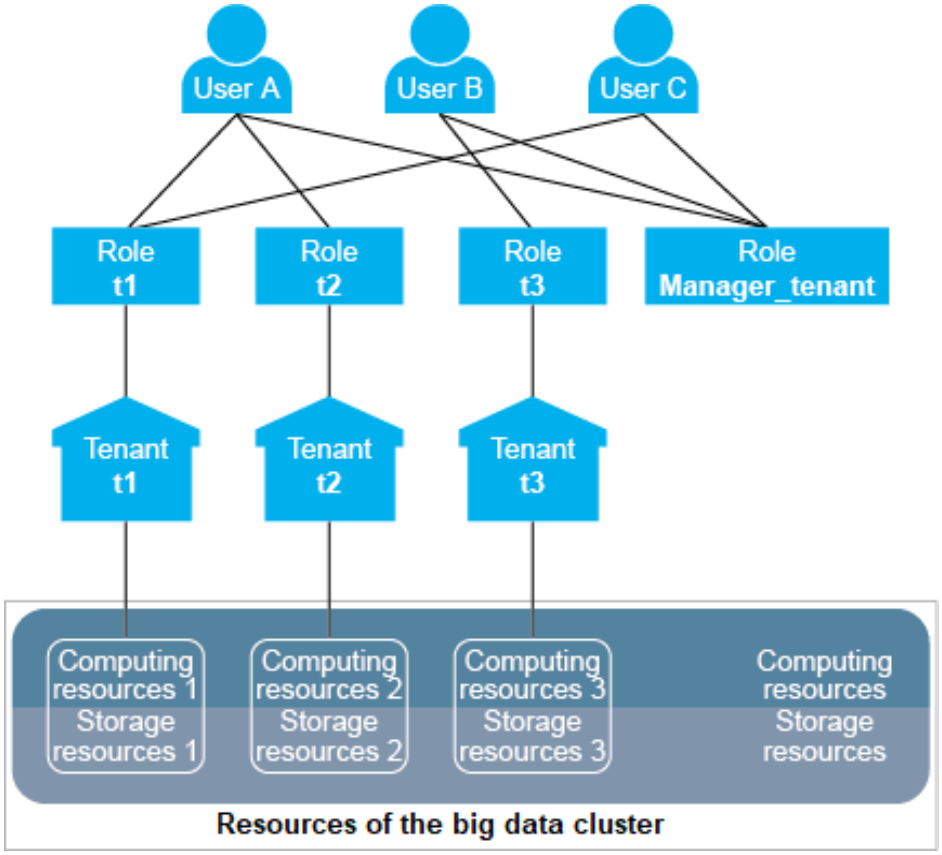
    In multi-tenant scenarios, the data of each tenant is stored separately to ensure data security. The access to tenants' resources is controlled to ensure access security.

Multi-tenant Model

- **Multi-tenant model**

  The following figure shows a multi-tenant model.

  Figure 8-7 Multi-tenant model



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-39 describes the concepts involved in **Figure 8-7**.

Table 8-39 Concepts in the model

Concept	Description
Users	A natural person who has a username and password and uses the big data cluster. There are three different users in the figure: users A, B, and C.
Role	A role is a carrier of one or more permissions. Permissions are assigned to specific objects, for example, access permissions for the **/tenant** directory in HDFS. There are four roles: **t1**, **t2**, **t3**, and **Manager_tenant**. • Roles **t1**, **t2**, and **t3** are automatically generated when tenants are created. The role names are the same as the tenant names. That is, roles **t1**, **t2**, and **t3** map to tenants **t1**, **t2**, and **t3**. Role names and tenant names need to be used in pair. • Role **Manager_tenant** is defaulted in the cluster and cannot be used separately.
Tenant	A tenant is a resource set in a big data cluster. Multiple tenants are referred to as multi-tenancy. The resource sets further divided under a tenant are called sub-tenants. There are three tenants: **t1**, **t2**, and **t3**.
Resource	• Compute resources include CPUs and memory. Compute resources are allocated to a tenant from the cluster's total resources. Each tenant's resources are isolated from others. According to the figure, compute resources 1, 2, and 3 are allocated to tenants **t1**, **t2**, and **t3** respectively from the cluster's total compute resources. • Storage resources include disks and third-party storage systems. Storage resources are allocated to each tenant from the cluster's total storage. Each tenant's storage is isolated from others. According to the figure, storage resources 1, 2, and 3 are allocated for tenants **t1**, **t2**, and **t3** respectively from the cluster's total storage resources.

To use a tenant's resources or add/delete sub-tenants, a user must be assigned both the tenant role and the **Manager_tenant** role. **Table 8-40** lists the roles assigned to each user in **Figure 8-7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-40 Roles assigned to each user

User	Role	Permission
User A	• Role **t1** • Role **t2** • Role **Manager_tenant**	• Uses the resources of tenants **t1** and **t2**. • Adds or deletes sub-tenants of tenants **t1** and **t2**.
User B	• Role **t3** • Role **Manager_tenant**	• Uses the resources of tenant **t3**. • Adds or deletes sub-tenants of tenant **t3**.
User C	• Role **t1** • Role **Manager_tenant**	• Uses the resources of tenant **t1**. • Adds or deletes sub-tenants of tenant **t1**.

A user can be assigned multiple roles, and one role can also be assigned to multiple users. Users are associated to tenants after being assigned the tenant roles. Therefore, tenants and users form a many-to-many relationship. Users can access resources across multiple tenants, and multiple users can share resources within the same tenant. For example, in **Figure 8-7**, user A uses the resources of tenants **t1** and **t2**, and users A and C uses the resources of tenant **t1**.

The concepts of a parent tenant, sub-tenant, level-1 tenant, and level-2 tenant are all designed for the multi-tenant service scenarios. Pay attention to the differences between these concepts and the concepts of a leaf tenant resource and non-leaf tenant resource on FusionInsight Manager.

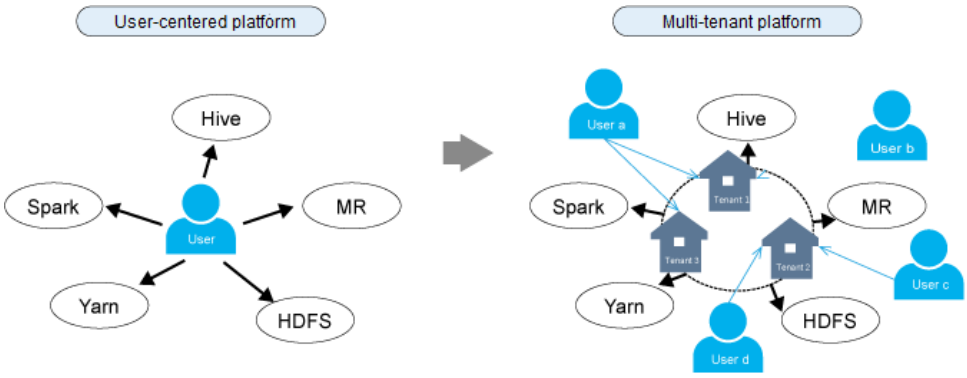
- **Level-1 tenant:** determined based on the tenant's level. For example, the first created tenant is a level-1 tenant and its sub-tenant is a level-2 tenant.
- **Parent tenant and sub-tenant:** indicates the hierarchical relationship between tenants.
- **Non-leaf tenant resource:** indicates the tenant type selected during tenant creation. This tenant type can be used to create sub-tenants.
- **Leaf tenant resource:** indicates the tenant type selected during tenant creation. This tenant type cannot be used to create sub-tenants.

- **Multi-tenant platform**:

  The MRS big data platform's core concept is the tenant, which enables the transition from user-centered to multi-tenant architecture to support enterprise multi-tenant applications. **Figure 8-8** shows the transformation of big data platforms.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 8-8 Platform transformation from user-centered to multi-tenant



On a user-centered big data platform, users can directly access and use all resources and services.

- However, user applications may use only partial cluster resources, resulting in low resource utilization.
- The data of different users may be stored together, decreasing data security.

On a multi-tenant big data platform, users use required resources and services by accessing the tenants.

- Resources are allocated and scheduled based on application requirements and used based on tenants, increasing resource utilization.
- Users can access the resources of tenants only after being associated with tenant roles, enhancing access security.
- The data of tenants is isolated, ensuring data security.

Multi-Tenant Resource

MRS cluster resources are classified into compute resources and storage resources. The multi-tenant architecture implements resource isolation.

- **Compute Resources**

  Compute resources include CPUs and memory. One tenant cannot occupy the compute resources of another tenant.

  Compute resources are divided into static service resources and dynamic resources.

  The resources allocated to YARN in a big data cluster are static service resources but can be dynamically allocated to job queues by YARN.

  - **Static Service Resources**

    Static service resources are compute resources allocated to each service and are not shared between services. The total compute resources of each service are fixed. These services include Flume, HBase, HDFS, and Yarn.

  - **Dynamic Resources**

    YARN provides distributed resource management for a big data cluster. The total volume of resources allocated to YARN can be configured. Then YARN allocates and schedules compute resources for job queues. For

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MapReduce, Spark, Flink, and Hive task queues, compute resources are allocated and scheduled by YARN.

YARN queues are fundamental units of scheduling compute resources.

The resources obtained by tenants using YARN queues are dynamic resources. Users can dynamically create and modify the queue quotas and view the status and statistics of the queues.

- **Resource pool:**

  Enterprise IT systems often face complex cluster environments and diverse upper-layer requirements, such as the following scenarios:

  - Heterogeneous cluster: The computing speed, storage capacity, and network performance of each node in the cluster are different. All the tasks of complex applications need to be properly allocated to each compute node in the cluster based on service requirements.
  - Computing isolation: Data must be shared among multiple departments but computing resources must be distributed onto different compute nodes.

  These require that the compute nodes be further partitioned.

  Resource pools are used to specify the configuration of dynamic resources. YARN queues are associated with resource pools for resource allocation and scheduling.

  One tenant can have only one default resource pool. Users can be assigned the role of a tenant to use the resources in the resource pool of the tenant. To use resources in multiple resource pools, a user can be assigned the roles of multiple tenants.

- **Dynamic resource scheduling:**

  YARN dynamic resources support label-based scheduling. This policy creates labels for compute nodes (YARN NodeManagers) and adds the compute nodes with the same label into the same resource pool. Then YARN dynamically associates the queues with resource pools based on the resource requirements of the queues.

  For example, a cluster has more than 40 nodes which are labeled by **Normal**, **HighCPU**, **HighMEM**, or **HighIO** based on their hardware and network configurations and added into four resource pools, respectively. **Table 8-41** describes the performance of each node in the resource pool.

  Table 8-41 Performance of each node in a resource pool

Label	Node Count	Hardware and Network Configuration	Added To	Associated With
Normal	10	Minor	Resource pool A	Common queue

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Label	Node Count	Hardware and Network Configuration	Added To	Associated With
HighCPU	10	High-performance CPU	Resource pool B	Computing-intensive queue
HighMEM	10	Large memory	Resource pool C	Memory-intensive queue
HighIO	10	High-performance network	Resource pool D	I/O-intensive queue

A queue can use only the compute nodes in its associated resource pool.

- A common queue is associated with resource pool A and uses **Normal** nodes with general hardware and network configurations.
- A computing-intensive queue is associated with resource pool B and uses **HighCPU** nodes with high-performance CPUs.
- A memory-intensive queue is associated with resource pool C and uses **HighMEM** nodes with large memory.
- An I/O-intensive queue is associated with resource pool D and uses **HighIO** nodes with high-performance network.

YARN queues are associated with specified resource pools to efficiently utilize resources in resource pools and maximize node performance.

FusionInsight Manager supports a maximum of 50 resource pools. The system has a default resource pool.

- **Storage resources**

  Storage resources include disks and third-party storage systems. One tenant cannot access the data of another tenant.

  As a distributed file storage service in a big data cluster, HDFS stores all the user data of the upper-layer applications in the big data cluster, including the data written to HBase tables or Hive tables.

  A directory is the basic unit of allocating HDFS storage resources. HDFS supports the conventional hierarchical file structure. Users or applications can create directories and create, delete, move, or rename files in directories. Tenants can obtain storage resources from specified directories in the HDFS file system.

  The storage resource scheduling mechanism is as follows:

  – HDFS directories can be stored on nodes with specified labels or disks of specified hardware types. For example, the service scenarios are as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

    - When both real-time query and data analysis tasks are running in the same cluster, the real-time query tasks need to be deployed only on certain nodes, and the task data must also be stored on these nodes.
    - Based on actual service requirements, key data needs to be stored on highly reliable nodes.
  - Administrators can flexibly configure HDFS data storage policies based on actual service requirements and data features to store data on specified nodes.
  - For tenants, storage resources refer to the HDFS resources they use. Data of specified directories can be stored to the tenant-specified storage paths, thereby implementing storage resource scheduling and ensuring data isolation between tenants.
  - Users can add or delete HDFS storage directories of tenants and set the file quantity quota and storage capacity quota of directories to manage storage resources.

Schedulers

Multi-tenant schedulers are classified into the open-source Capacity scheduler and enhanced Superior scheduler. By default, the Superior scheduler is enabled for the MRS cluster.

- The Capacity scheduler is an open-source scheduler.
- The Superior scheduler is an enhanced version and named after the Lake Superior, indicating that the scheduler can manage a large amount of data.

NOTE

You can check the scheduler type of YARN from the **yarn.resourcemanager.scheduler.class** value. For details about how to switch the scheduler, see **Switching the MRS Tenant Resource Scheduler**.

To meet enterprise requirements and tackle scheduling challenges faced by the YARN community, the Superior scheduler makes the following enhancements in addition to inheriting the advantages of the Capacity scheduler and Fair scheduler:

- Enhanced resource sharing policy

  The Superior scheduler supports queue hierarchy. It integrates the functions of open-source schedulers and shares resources based on configurable policies. In terms of instances, administrators can use the Superior scheduler to configure an absolute value or percentage policy for queue resources. The resource sharing policy of the Superior scheduler enhances label-based scheduling of YARN as a resource pool feature. The nodes in the YARN cluster can be grouped based on the capacity or service type to ensure that queues can more efficiently utilize resources.
- Tenant-based resource reservation policy

  Some tenants may run critical tasks at some time, and their resource requirements must be preferentially addressed. The Superior scheduler builds a mechanism to support the resource reservation policy. Reserved resources can be allocated to the critical tasks running in the specified tenant queues in a timely manner to ensure proper task execution.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Fair sharing among tenants and resource pool users

  The Superior scheduler allows shared resources to be configured for users in a queue. Each tenant may have users with different weights. Heavily weighted users may require more shared resources.

- Ensured scheduling performance in a big cluster

  The Superior scheduler receives heartbeats from each NodeManager and saves resource information in memory, which enables the scheduler to control cluster resource usage globally. The Superior scheduler uses the push scheduling model, which makes the scheduling more precise and efficient and remarkably improves cluster resource utilization. Additionally, the Superior scheduler delivers excellent performance when the interval between NodeManager heartbeats is long and prevents heartbeat storms in big clusters.

- Priority policy

  If the minimum resource requirement of a service cannot be met after the service obtains all available resources, a preemption occurs. The preemption function is disabled by default.

Multi-Tenant Management

- **Unified multi-tenant management**

  Log in to the MRS console or FusionInsight Manager and click **Tenant Resources**. The displayed page integrates multiple functions such as tenant lifecycle management, tenant resource configuration, tenant service association, and tenant resource usage statistics, delivering a mature multi-tenant management model and achieving centralized tenant and service management.

  - Intuitive UI: MRS provides the graphical multi-tenant management interface and manages and operates multiple levels of tenants using the tree structure. Additionally, it integrates the basic information and resource quota of the current tenant in one interface to facilitate O&M and management.
  - Hierarchical tenant management: MRS allows you to add sub-tenants to an existing tenant to re-configure resources. Sub-tenants of level-1 tenants are level-2 tenants. So on and so forth. FusionInsight Manager provides enterprises with a field-tested multi-tenant management model, enabling centralized tenant and service management.

- **Simplified permission management**

  MRS hides internal permission management details from common users and simplifies permission management operations for administrators, improving usability and user experience of tenant permission management.

  - MRS employs role-based access control (RBAC) to configure different permissions for users based on service scenarios during multi-tenant management.
  - The administrator of tenants has tenant management permissions, including viewing resources and services of the current tenant, adding or deleting sub-tenants of the current tenant, and managing permissions of sub-tenants' resources. MRS supports setting of the administrator for a single tenant so that the management over this tenant can be delegated to a user who is not the system administrator.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Roles of a tenant have all permissions on the computing resources and storage resources of the tenant. When a tenant is created, the system automatically creates roles for this tenant. You can add a user and assign the tenant roles to the user so that the user can use the resources of the tenant.

- **Easy resource management**
  - Free resource configuration

    You can configure the compute resources and storage resources during the creation of a tenant and add, modify, or delete the resources of the tenant.

    Permissions of the roles that are granted to a tenant are updated automatically when you modify the computing or storage resources of the tenant.
  - Resource usage statistics

    Resource usage statistics are critical for administrators to determine O&M activities based on the status of cluster applications and services, improving the cluster O&M efficiency. MRS displays the resource statistics of tenants in **Resource Quota**, including the vCores, memory, and HDFS storage resources.

    - The available resources of the Capacity scheduler and Superior scheduler are calculated as follows:
      - Capacity

        Available YARN resources (memory and CPU) = Resource capacity (%) x Total capacity of the resource pool

        When a queue spans multiple resource pools, its available resources are the combined total of resources allocated by each pool.
      - Superior

        The available YARN resources (memory and CPU) are allocated in proportion based on the queue weight.
    - When the tenant administrator is bound to a tenant role, the tenant administrator has the permissions to manage the tenant and use all resources of the tenant.
  - Graphical resource monitoring

    Graphical resource monitoring supports the graphical display of monitoring metrics listed in **Table 8-42**.

    By default, real-time monitoring data is displayed. You can click to customize a time range. Click and choose **Export** from the shortcut menu to export monitoring information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-42 Monitoring metrics

Service	Metric Item	Description
HDFS	HDFS Tenant Space Details • Allocated Space • Used Space	HDFS can monitor a specified storage directory. The storage directory is the same as the directory added by the current tenant in **Resource**.
	HDFS Tenant File Object Details • Number of Used File Objects	
YARN	Yarn Allocated Cores • Maximum Number of CPU Cores in an AM • Allocated Cores • Number of Used CPU Cores in an AM	Monitoring information of the current tenant is displayed. If no sub-item is configured for a tenant, this information is not displayed. The monitoring data is obtained from **Scheduler** > **Application Queues** > **Queue:** *Tenant name* on the native web UI of YARN.
	Yarn Allocated Memory • Allocated Maximum AM Memory • Allocated Memory • Used AM Memory	

8.9.2 Using MRS Multi-Tenancy

Scenarios

Tenants are used in resource control and service isolation scenarios. You need to determine the service scenarios of cluster resources and then plan tenants.

Multi-tenancy involves three types of operations: creating a tenant, managing tenants, and managing resources. **Table 8-43** describes these operations.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-43 Multi-tenant operations

Operation	Action	Description
Creating a tenant	• Creating a tenant • Creating a sub-tenant • Adding a user and assigning a tenant role to the user	During the creation of a tenant, you can configure its computing resources, storage resources, and associated services based on service requirements. In addition, you can add users to the tenant and bind necessary roles to these users. A user to create a level-1 tenant needs to be bound to the **Manager_administrator** or **System_administrator** role. A user to create a sub-tenant needs to be bound to the role of the parent tenant at least.
Managing tenants	• Managing a tenant directory • Restoring tenant data • Clearing non-associated queues of a tenant • Deleting a tenant	You can edit tenants as services change. A user to manage or delete a level-1 tenant or restore tenant data needs to be assigned the **Manager_administrator** or **System_administrator** role. A user to manage or delete a sub-tenant needs to be assigned the role of the parent tenant at least.
Managing resources	• Creating a resource pool • Modifying a resource pool • Deleting a resource pool • Configuring a queue • Configuring the queue capacity policy of a resource pool • Clearing configuration of a queue	You can reconfigure resources for tenants as the services change. A user to manage resources needs to be assigned the **Manager_administrator** or **System_administrator** role.

Creating Tenants and Sub-Tenants

Administrators need to determine the service scenarios of cluster resources and then plan tenants. After that, administrators add tenants and configure dynamic resources, storage resources, and associated services for the tenants on MRS.

Table 8-44 shows the process for creating a tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 8-9 Creating a tenant

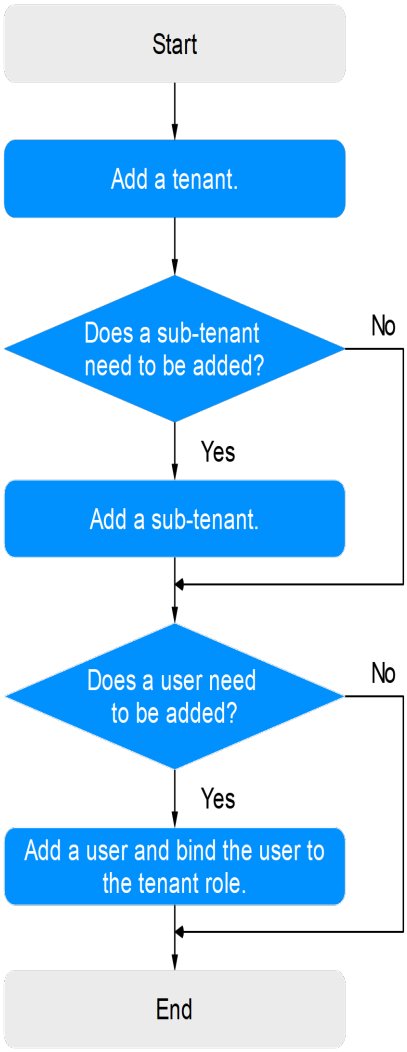


Table 8-44 Operations for creating a tenant

Operation	Description
Creating a tenant	You can configure the compute resources, storage resources, and associated services of the tenant.
Creating a sub-tenant	You can configure the compute resources, storage resources, and associated services of the sub-tenant.
Adding a user and assigning the tenant role to the user	If a user wants to use the resources of tenant **tenant1** or add or delete sub-tenants for **tenant1**, the user must be bound to both the **Manager_tenant** and **tenant1**_*Cluster ID* roles.

8.9.3 Configuring MRS Tenants

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.9.3.1 Creating an MRS Tenant

You can create a tenant on MRS Manager to specify the resource usage.

Prerequisites

- A tenant name has been planned. The name must not be the same as that of a role or YARN queue that exists in the current cluster.
- If a tenant requires storage resources, a storage directory has been planned based on service requirements, and the planned directory does not exist under the HDFS directory.
- The resources that can be allocated to the current tenant have been planned and the sum of the resource percentages of direct sub-tenants/sub-tenants under the parent tenant at every level does not exceed 100%.
- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Adding a Tenant on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the MRS cluster details page, click **Tenant**.

Step 4 Click **Create Tenant**. On the displayed page, configure tenant properties by referring to the following table.

Table 8-45 Tenant property parameters (MRS 3.x)

Parameter	Example Value	Description
Name	ta1	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. In some versions, this parameter is named **Tenant Resource Type**. ● When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. ● When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant. If you select **ClickHouse** or **Doris** for **Service**, this parameter can only be set to **Leaf Tenant**.
Compute Resource	-	Dynamic compute resources for the current tenant. ● When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. ● If **Yarn** is not selected, the system does not automatically create a queue.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Configuration Mode	-	If **Yarn** is selected for **Compute Resource**, this parameter can be set to **Basic** or **Advanced**. ● **Basic**: Configure the percentage of compute resources used by the tenant in the default resource pool by specifying **Default Resource Pool Capacity (%)**. ● **Advanced**: Configure the following parameters for advanced settings: – **Weight**: Tenant resource weight. The value ranges from 0 to 100. Tenant resource weight = Tenant weight/Total weight of tenants at the same level – **Minimum Resources**: resources preempted by the tenant. The value is a percentage or absolute value of the parent tenant's resources. When a tenant's workload is light, their resources are automatically lent to other tenants. When available resources are fewer than **Minimum Resources**, the tenant can preempt the resources that were lent out. – **Maximum Resources**: Maximum resources that can be used by the tenant. The value is a percentage or absolute value of the parent tenant's resources. – **Reserved Resources**: Resources reserved for the tenant. The reserved resources cannot be used by other tenants even if no job is using resources of the tenant. The value is a percentage or absolute value of the parent tenant's resources.
Storage Resource	HDFS	Storage resources selected for the current tenant. ● If **HDFS** is selected, the system automatically creates a folder named after the tenant name in the **/tenant** directory. When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. ● If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS.
Maximum Number of Files/ Directories	10	Maximum number of files or directories that can be created in HDFS. You need to configure this parameter only when **Storage Resource** is set to **HDFS**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Space Quota	500 MB	Specifies the quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. ● The minimum value is **1**, and the unit is MB or GB. ● This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. ● If the value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two replicas are generated in total. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1	HDFS directory for tenant resource data. ● The system automatically creates a file folder named after the tenant name in the **/tenant** directory by default. For example, the default HDFS storage directory for tenant **ta1** is **/tenant/ta1**. ● When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. The storage path is customizable.
Service	-	Other service resources associated with the tenant. (This parameter is grayed out if the cluster has no components that support service association.) Click **Associate Service** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **share** is selected, service resources are shared.
Description	-	Specifies the description of the current tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-46 Tenant parameters (MRS 2.x and earlier versions)

Parameter	Example Value	Description
Name	ta1	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant.
Compute Resource	-	Dynamic compute resources for the current tenant. (In some versions, the parameter is named **Dynamic Resource**.) • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.
Default Resource Pool Capacity (%)	80	Percentage of the compute resources used by the current tenant in the **default** resource pool. You need to configure this parameter only when **Computing Resource** is set to **Yarn**.
Default Resource Pool Max. Capacity (%)	70	Maximum percentage of the compute resources used by the current tenant in the **default** resource pool. You need to configure this parameter only when **Computing Resource** is set to **Yarn**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Resource	HDFS	Storage resources selected for the current tenant. • If **HDFS** is selected, the system automatically creates a folder named after the tenant name in the **/tenant** directory. When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. • If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS.
Storage Space Quota (MB)	500 MB	Specifies the quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. • The minimum value is **1**, and the unit is MB or GB. • This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. • If the value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two replicas are generated in total. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1	HDFS storage directory for the tenant. • The system automatically creates a file folder named after the tenant name in the **/tenant** directory by default. For example, the default HDFS storage directory for tenant **ta1** is **/tenant/ta1**. • When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. The storage path is customizable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Service	-	Other service resources associated with the tenant. (This parameter is grayed out if the cluster has no components that support service association.) Click **Associate Service** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **share** is selected, service resources are shared.
Description	-	Specifies the description of the current tenant.

Step 5 Click **OK** to save. Wait until the tenant is created.

In the tenant list on the left, click the name of the created tenant to view its basic information, resource quotas, and charts.

- Roles, compute resources, and storage resources are automatically created when tenants are created.
- The new role has permissions on the compute and storage resources. The role and its permissions are controlled by the system automatically and cannot be controlled manually under **Manage Role**.
- If you want to use the tenant, create a system user and assign the Manager_tenant role and the role corresponding to the tenant to the user. For details, see **Binding a Tenant to an MRS Cluster User**.

----End

Adding a Tenant on Manager (for MRS 3.x and Later)

Step 1 Log in to MRS Manager of the cluster and choose **Tenant Resources**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Click ⊕. On the page that is displayed, configure tenant properties according to **Table 8-47**.

If the MRS cluster is a 3.x version earlier than 3.3.0, select the cluster for which you want to create a tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-47 Tenant property parameters (MRS 3.x)

Parameter	Example Value	Description
Name	ta1	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. In some versions, this parameter is named **Tenant Resource Type**. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant. If you select **ClickHouse** or **Doris** for **Service**, this parameter can only be set to **Leaf Tenant**.
Compute Resource	-	Dynamic compute resources for the current tenant. • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Configuration Mode	-	If **Yarn** is selected for **Compute Resource**, this parameter can be set to **Basic** or **Advanced**. This parameter must be set only for the Superior scheduler. • **Basic**: Configure the percentage of compute resources used by the tenant in the default resource pool by specifying **Default Resource Pool Capacity (%)**. • **Advanced**: Configure the following parameters for advanced settings: – **Weight**: Tenant resource weight. The value ranges from 0 to 100. Tenant resource weight = Tenant weight/Total weight of tenants at the same level – **Minimum Resources**: resources preempted by the tenant. The value is a percentage or absolute value of the parent tenant's resources. When a tenant's workload is light, their resources are automatically lent to other tenants. When available resources are fewer than **Minimum Resources**, the tenant can preempt the resources that were lent out. – **Maximum Resources**: maximum resources that can be used by a tenant. The value is a percentage or absolute value of the parent tenant's resources. – **Reserved Resources**: resources reserved for the tenant. The reserved resources cannot be used by other tenants even if no job is using resources of the tenant. The value is a percentage or absolute value of the parent tenant's resources.
Default Resource Pool Capacity (%)	20	Percentage of compute resources used by the current tenant in the default resource pool. The value ranges from 0 to 100%.
Default Resource Pool Max Capacity (%)	100	Maximum percentage of compute resources that can be used by the current tenant in the default resource pool. The value ranges from 0 to 100%. This parameter must be set only for the Capacity scheduler.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Resource	HDFS	Storage resources selected for the current tenant. ● If **HDFS** is selected, the system automatically creates a folder named after the tenant name in the **/tenant** directory. When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. ● If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS.
Maximum Number of Files/ Directories	10	Maximum number of files or directories that can be created in HDFS. You need to configure this parameter only when **Storage Resource** is set to **HDFS**.
Storage Space Quota	500 MB	Specifies the quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. ● The minimum value is **1**, and the unit is MB or GB. ● This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. ● If the value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two replicas are generated in total. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1	HDFS directory for tenant resource data. ● The system automatically creates a folder named after the tenant name in the **/tenant** directory by default. For example, the default HDFS storage directory for tenant **ta1** is **/tenant/ta1**. ● When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. The storage path is customizable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Service	-	Other service resources associated with the tenant. (This parameter is grayed out if the cluster has no components that support service association.) Click **Associate Service** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **share** is selected, service resources are shared.
Description	-	Specifies the description of the current tenant.

Roles, compute resources, and storage resources are automatically created when tenants are created.

- The new role has permissions on the compute and storage resources. This role and its permissions are automatically controlled by the system and cannot be manually managed by choosing **System** > **Permission** > **Role**. The role name is in the format of *Tenant name_Cluster ID*. The ID of the first cluster is not displayed by default.
- When using this tenant, create a system user and assign the tenant role to the user. For details, see **Binding a Tenant to an MRS Cluster User**.
- During the tenant creation, the system automatically creates a YARN queue named after the tenant. If the queue name already exists, the new queue is named *Tenant name*-*N*. *N* indicates a natural number starting from **1**. When a same name exists, the value *N* increases automatically to differentiate the queue from others. For example, **saletenant**, **saletenant-1**, and **saletenant-2**.

If you need to associate a service, click **Associate Service**, configure other service resources associated with the tenant by referring to the following instructions, and click **OK**.

- Set **Service** to **HBase** and **Association Type** to **Exclusive** or **Shared**.
  - **Exclusive** indicates that the service resources are used by the tenant exclusively and cannot be associated with other tenants.
  - **Shared** indicates that the service resources can be shared with other tenants.
- MRS 3.2.0 or later: Select **ClickHouse** for **Service**.
  - **Association Type**: When **Service** is set to **ClickHouse**, **Association Type** can only be set to **Shared**. MRS 3.3.0 and later versions support **Exclusive** and **Shared**.
  - **Associate Logical Cluster**: If the logical cluster function is not enabled for ClickHouse, **default_cluster** is selected by default. If the function is enabled, select the logical cluster to which you want to associate.
  - **CPU Priority**: The CPU priority ranges from –20 to 19. This value is associated with the NICE value of the OS. A smaller value indicates a higher CPU priority.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - **Memory**: The maximum value of this parameter is **100**, in percentage. For example, if this parameter is set to **80**, the total memory that can be used by the current tenant is calculated as follows: Available memory x 80%.
  - **Concurrency**: The maximum number of concurrent resources available for all the users bound to the tenant. This parameter is required for clusters of MRS 3.3.0 or later.
- MRS 3.5.0 or later: Select **Doris** for **Service**.
  - **Association Type**: The default value is **Shared**.
  - **CPU Quota Usage**: CPU usage per unit time. The value ranges from 1 to 100.
  - **Memory Quota**: proportion of available memory. For example, if the value is **20**, the available memory of the current tenant on each BE instance node is: Physical memory x **mem_limit** (BE memory limit in percentage) x 20%.
  - **Concurrency**: maximum number of concurrent query tasks that can be executed on a single FE instance.
  - **Queue Length**: maximum number of query tasks waiting to be executed in the queue because the number of concurrent query tasks exceeds the upper limit.
  - **Waiting Duration (ms)**: maximum waiting time of a task in the queue, in milliseconds
  - **Soft Memory Limit**: whether a tenant can use more memory resources than it is allowed If this function is enabled and the system has idle memory resources, you can use more memory than the limit. Your tasks that occupy the most memory in the group will be canceled only when resources are insufficient.

Only HBase can be associated with a new tenant. However, HDFS, HBase, and YARN can be associated with existing tenants.

- To associate an existing tenant with service resources, click the target tenant in the tenant list, switch to the **Service Associations** page, and click **Associate Service** to configure resources to be associated with the tenant.
- To disassociate an existing tenant from service resources, click the target tenant in the tenant list, switch to the **Service Associations** page, and click **Delete** in the **Operation** column. In the displayed dialog box, select **I have read the information and understand the impact** and click **OK**.

Step 3 Click **OK**. Wait until the system displays a message indicating that the tenant is successfully created.

For clusters of MRS 3.3.0 or later, click to export basic information about all tenants in the current cluster.

----End

Adding a Tenant on Manager (for MRS 2.x and Earlier)

Step 1 Log in to MRS Manager of the cluster and choose **Tenant**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Click **Create Tenant**. On the displayed page, configure tenant properties.

Table 8-48 Tenant parameters (MRS 2.x and earlier versions)

Parameter	Example Value	Description
Name	ta1	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant.
Compute Resource	-	Dynamic compute resources for the current tenant. (In some versions, the parameter is named **Dynamic Resource**.) • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.
Default Resource Pool Capacity (%)	80	Percentage of the compute resources used by the current tenant in the **default** resource pool. You need to configure this parameter only when **Computing Resource** is set to **Yarn**.
Default Resource Pool Max. Capacity (%)	70	Maximum percentage of the compute resources used by the current tenant in the **default** resource pool. You need to configure this parameter only when **Computing Resource** is set to **Yarn**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Resource	HDFS	Storage resources selected for the current tenant. • If **HDFS** is selected, the system automatically creates a folder named after the tenant name in the **/tenant** directory. When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. • If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS.
Storage Space Quota (MB)	500 MB	Specifies the quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. • The minimum value is **1**, and the unit is MB or GB. • This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. • If the value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two replicas are generated in total. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1	HDFS storage directory for the tenant. • The system automatically creates a folder named after the tenant name in the **/tenant** directory by default. For example, the default HDFS storage directory for tenant **ta1** is **/tenant/ta1**. • When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. The storage path is customizable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Service	-	Other service resources associated with the tenant. (This parameter is grayed out if the cluster has no components that support service association.) Click **Associate Service** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **share** is selected, service resources are shared.
Description	-	Specifies the description of the current tenant.

Step 3 Click **OK** to save. Wait until the tenant is created.

- Roles, compute resources, and storage resources are automatically created when tenants are created.
- The new role has permissions on the compute and storage resources. The role and its permissions are controlled by the system automatically and cannot be controlled manually under **Manage Role**.
- If you want to use the tenant, create a system user and assign the Manager_tenant role and the role corresponding to the tenant to the user. For details, see **Binding a Tenant to an MRS Cluster User**.

----End

8.9.3.2 Creating an MRS Sub-Tenant

You can create sub-tenants on MRS and allocate resources of the current tenant to the sub-tenants based on the resource consumption and isolation planning and requirements of services.

If the tenant is a non-leaf tenant, you can add sub-tenants. Sub-tenants cannot be created for a leaf tenant.

Prerequisites

- A non-leaf tenant has been added by referring to **Creating an MRS Tenant**.
- A tenant name has been planned. The name must not be the same as that of a role or YARN queue that exists in the current cluster.
- If a sub-tenant requires storage resources, a storage directory has been planned based on service requirements, and the planned directory does not exist under the storage directory of the parent tenant.
- The resources that can be allocated to the current tenant have been planned and the sum of the resource percentages of direct sub-tenants under the parent tenant at every level does not exceed 100%.
- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Adding an MRS Sub-Tenant on the MRS Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Tenants**.

Step 4 In the tenant list on the left, move the cursor to the tenant node to which a sub-tenant is to be added. Click **Create sub-tenant**. On the displayed page, configure the sub-tenant attributes according to the following table:

Table 8-49 Tenant parameters (MRS 3.x)

Parameter	Example Value	Description
Name	ta1	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf	Whether the tenant is a leaf tenant. In some versions, this parameter is named **Tenant Resource Type**. • When **Leaf** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant. If you select **ClickHouse** or **Doris** for **Service**, this parameter can only be set to **Leaf Tenant**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Compute Resource	-	Dynamic compute resources for the current tenant. ● When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. ● If **Yarn** is not selected, the system does not automatically create a queue.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Configuration Mode	-	If **Yarn** is selected for **Compute Resource**, this parameter can be set to **Basic** or **Advanced**. ● **Basic**: Configure the percentage of compute resources used by the tenant in the default resource pool by specifying **Default Resource Pool Capacity (%)**. ● **Advanced**: Configure the following parameters for advanced settings: – **Weight**: Tenant resource weight. The value ranges from 0 to 100. Tenant resource weight = Tenant weight/Total weight of tenants at the same level – **Minimum Resources**: Guaranteed resource allocation for the tenant (preemption supported). The value is a percentage or absolute value of the parent tenant's resources. When a tenant's workload is light, their resources are automatically lent to other tenants. When available resources are fewer than **Minimum Resources**, the tenant can preempt the resources that were lent out. – **Maximum Resources**: Maximum resources that can be used by the tenant. The value is a percentage or absolute value of the parent tenant's resources. – **Reserved Resources**: Resources reserved for the tenant. The reserved resources cannot be used by other tenants even if no job is using resources of the tenant. The value is a percentage or absolute value of the parent tenant's resources.
Storage Resource	HDFS	Storage resources selected for the current tenant. ● If **HDFS** is selected, the system automatically creates a folder named after the tenant name in the **/tenant** directory. When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. ● If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS.
Maximum Number of Files/ Directories	10	Maximum number of files or directories that can be created in HDFS. You need to configure this parameter only when **Storage Resource** is set to **HDFS**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Space Quota	500 MB	Quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. ● The minimum value is **1**, and the unit is MB or GB. ● This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. ● If its value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. To ensure data reliability, the system automatically generates one backup file when a file is stored in the HDFS. That is, two replicas of the same file are stored by default. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1	HDFS directory for tenant resource data. ● The system automatically creates a folder named after the tenant name in the **/tenant** directory by default. For example, the default HDFS storage directory for tenant **ta1** is **/tenant/ta1**. ● When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. The storage path is customizable.
Service	-	Other service resources associated with the tenant. (This parameter is grayed out if the cluster has no components that support service association.) Click **Associate** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **Shared** is selected, service resources are shared.
Description	-	Description of the tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-50 Tenant parameters (MRS 2.x and earlier versions)

Parameter	Example Value	Description
Name	ta1	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant.
Compute Resource	-	Dynamic compute resources for the current tenant. (In some versions, the parameter is named **Dynamic Resource**.) • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.
Default Resource Pool Capacity (%)	80	Percentage of the compute resources used by the current tenant in the **default** resource pool. You need to configure this parameter only when **Computing Resource** is set to **Yarn**.
Default Resource Pool Max. Capacity (%)	70	Maximum percentage of the compute resources used by the current tenant in the **default** resource pool. You need to configure this parameter only when **Computing Resource** is set to **Yarn**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Resource	HDFS	Storage resources selected for the current tenant. • If **HDFS** is selected, the system automatically creates a folder named after the tenant name in the **/tenant** directory. When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. • If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS.
Storage Space Quota (MB)	500 MB	Quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. • The minimum value is **1**, and the unit is MB or GB. • This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. • If its value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. To ensure data reliability, the system automatically generates one backup file when a file is stored in the HDFS. That is, two replicas of the same file are stored by default. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1	HDFS storage directory for the tenant. • The system automatically creates a folder named after the tenant name in the **/tenant** directory by default. For example, the default HDFS storage directory for tenant **ta1** is **/tenant/ta1**. • When a tenant is created for the first time, the system automatically creates the **/tenant** directory in the HDFS root directory. The storage path is customizable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Service	-	Other service resources associated with the tenant. (This parameter is grayed out if the cluster has no components that support service association.) Click **Associate Service** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **Shared** is selected, service resources are shared.
Description	-	Description of the tenant.

Step 5 Click **OK** to save the settings.

It takes a few minutes to save the settings. If the **Tenant created successfully** is displayed in the upper-right corner, the tenant is added successfully. The tenant is created successfully.

- Roles, compute resources, and storage resources are automatically created when tenants are created.
- The new role has permissions on the compute and storage resources. This role and its permissions are automatically controlled by the system and cannot be manually managed by choosing **System** > **Permission** > **Role**. The role name is in the format of *Tenant name_Cluster ID*. The ID of the first cluster is not displayed by default.
- When using this tenant, create a system user and assign the tenant role to the user. For details, see **Binding a Tenant to an MRS Cluster User**.
- The sub-tenant can further allocate the resources of its parent tenant. The sum of the resource percentages of direct sub-tenants under a parent tenant at each level cannot exceed 100%. The sum of the compute resource percentages of all level-1 tenants cannot exceed 100%.

----End

Adding a Sub-Tenant on Manager (for MRS 3.x and Later)

Step 1 Log in to MRS Manager of the cluster and choose **Tenant Resources**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 In the tenant list on the left, select a parent tenant and click . On the displayed page, set the parameters.

- To use a Superior scheduler cluster, configure sub-tenant properties by referring to **Table 8-51**.
- To use the Capacity scheduler cluster, configure sub-tenant properties by referring to **Table 8-52**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- YARN in a new cluster uses the Superior scheduler by default. You can also switch the scheduler by referring to **Switching the MRS Tenant Resource Scheduler**.
- To query the scheduler type, log in to Manager and search for the **yarn.resourcemanager.scheduler.class** parameter on the **All Configurations** page of the YARN service.

Table 8-51 Sub-tenant parameters (Superior scheduler)

Parameter	Example Value	Description
Cluster	MRS Cluster	Cluster to which the parent tenant belongs.
Parent Tenant Resource	ta1	Name of the parent tenant. (In some versions, the parameter is named **Parent Tenant**.)
Name	ta1s	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Resource Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. (In some versions, the parameter is named **Tenant Type**.) • When **Leaf Tenant Resource** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. (In some versions, the parameter is named **Leaf Tenant**.) • When **Non-leaf Tenant Resource** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant. However, the tenant depth cannot exceed 5 levels. (In some versions, the parameter is named **Non-Leaf Tenant**.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Computing Resource	Yarn	Dynamic compute resources for the current tenant. • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Configuration Mode	-	Configuration mode of compute resource parameters If **Yarn** is selected for **Compute Resource**, this parameter can be set to **Basic** or **Advanced**. • **Basic**: Configure the percentage of compute resources used by the tenant in the default resource pool by specifying **Default Resource Pool Capacity (%)**. • **Advanced**: Configure the following parameters for advanced settings: – **Weight**: Tenant resource weight. The value ranges from 0 to 100. Tenant resource weight = Tenant weight/Total weight of tenants at the same level – **Minimum Resources**: resources preempted by the tenant. The value is a percentage or absolute value of the parent tenant's resources. When a tenant's workload is light, their resources are automatically lent to other tenants. When available resources are fewer than **Minimum Resources**, the tenant can preempt the resources that were lent out. – **Maximum Resources**: Maximum resources that can be used by the tenant. The value is a percentage or absolute value of the parent tenant's resources. – **Reserved Resources**: Resources reserved for the tenant. The reserved resources cannot be used by other tenants even if no job is using resources of the tenant. The value is a percentage or absolute value of the parent tenant's resources.
Storage Resource	HDFS	Storage resources selected for the current tenant. • When **HDFS** is selected, the system automatically creates a folder named after the sub-tenant in the HDFS parent tenant directory. • If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS. If the parent tenant does not have storage resources, the sub-tenant cannot use storage resources.
Quota	100	Quota for files and directories. You need to configure this parameter only when **Storage Resource** is set to **HDFS**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Space Quota	500 MB	Quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. • The minimum value is **1**, and the maximum value is the total storage quota of the parent tenant. The unit is MB or GB. • This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. • If its value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. • If this quota is greater than the quota of the parent tenant, the actual storage space does not exceed the quota of the parent tenant. To ensure data reliability, the system automatically generates one backup file when a file is stored in the HDFS. That is, two replicas of the same file are stored by default. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1/ta1s	HDFS storage directory for the tenant. • The system automatically creates a folder named after the sub-tenant name in the directory of the parent tenant by default. For example, if the sub-tenant is **ta1s** and the parent directory is **tenant/ta1**, the system sets this parameter for the sub-tenant to **tenant/ta1/ta1s**. • The storage path is customizable in the parent directory. The parent directory for the storage path must be the storage directory of the parent tenant.
Service	-	Whether to associate resources of other services. For details, see **Step 4**.
Description	-	Description of the tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-52 Sub-tenant parameters (Capacity scheduler)

Parameter	Example Value	Description
Cluster	MRS Cluster	Cluster to which the parent tenant belongs.
Parent Tenant Resource	ta1	Name of the parent tenant. (In some versions, the parameter is named **Parent Tenant**.)
Name	ta1s	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant.
Computing Resource	Yarn	Dynamic compute resources for the current tenant. • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.
Default Resource Pool Capacity (%)	-	Percentage of compute resources used by the current tenant. The base value is the total resources of the parent tenant.
Default Resource Pool Max Capacity (%)	90	Specifies the maximum percentage of the compute resources used by the current tenant. The base value is the total resources of the parent tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Resource	HDFS	Storage resources selected for the current tenant. ● When **HDFS** is selected, the system automatically creates a folder named after the sub-tenant in the HDFS parent tenant directory. ● If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS. If the parent tenant does not have storage resources, the sub-tenant cannot use storage resources.
Quota	100	Quota for files and directories. You need to configure this parameter only when **Storage Resource** is set to **HDFS**.
Storage Space Quota	-	Quota for HDFS storage space used by the current tenant. You need to configure this parameter only when **Storage Resource** is set to **HDFS**. ● The minimum value is **1**, and the maximum value is the total storage quota of the parent tenant. The unit is MB or GB. ● This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. ● If its value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. ● If this quota is greater than the quota of the parent tenant, the actual storage space does not exceed the quota of the parent tenant. To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two replicas are generated in total. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Storage Path	tenant/ta1/ta1s	HDFS storage directory for the tenant. • The system automatically creates a folder named after the sub-tenant name in the directory of the parent tenant by default. For example, if the sub-tenant is **ta1s** and the parent directory is **tenant/ta1**, the system sets this parameter for the sub-tenant to **tenant/ta1/ta1s**. • The storage path is customizable in the parent directory. The parent directory for the storage path must be the storage directory of the parent tenant.
Description	-	Description of the current tenant.

- Roles, compute resources, and storage resources are automatically created when tenants are created.
- The new role has permissions on the compute and storage resources. This role and its permissions are automatically controlled by the system and cannot be manually managed by choosing **System** > **Permission** > **Role**. The role name is in the format of *Tenant name_Cluster ID*. The ID of the first cluster is not displayed by default.
- When using this tenant, create a system user and assign the tenant role to the user. For details, see **Binding a Tenant to an MRS Cluster User**.
- The sub-tenant can further allocate the resources of its parent tenant. The sum of the resource percentages of direct sub-tenants under a parent tenant at each level cannot exceed 100%. The sum of the compute resource percentages of all level-1 tenants cannot exceed 100%.

Step 3 Check whether the current tenant needs to be associated with resources of other services.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Click **Associate Service** to configure other service resources used by the current tenant.

1. Select the target service from the **Service** drop-down list.
2. Set **Association Type** as follows:
   - **Exclusive** indicates that the service resources are used by the tenant exclusively and cannot be associated with other tenants.
   - **Shared** indicates that the service resources can be shared with other tenants.

Only HBase can be associated with a new tenant. However, HDFS, HBase, and YARN can be associated with existing tenants.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- To associate an existing tenant with service resources, click the target tenant in the tenant list, switch to the **Service Associations** page, and click **Service Association** to configure resources to be associated with the tenant.
- To disassociate an existing tenant from service resources, click the target tenant in the tenant list, switch to the **Service Associations** page, and click **Delete** in the **Operation** column. In the displayed dialog box, select **I have read the information and understand the impact**. and click **OK**.

3. Click **OK**.

Step 5 Click **OK**. Wait until the system displays a message indicating that the tenant is successfully created.

----End

Adding a Sub-Tenant on MRS Manager (for MRS 2.x and Earlier)

Step 1 Log in to MRS Manager of the cluster and choose **Tenant**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 In the tenant list on the left, move the cursor to the tenant node to which a sub-tenant is to be added. Click **Create sub-tenant**. On the displayed page, configure the sub-tenant attributes according to the following table:

Table 8-53 Sub-tenant parameters

Parameter	Example Value	Description
Parent tenant	ta1	Name of the parent tenant.
Name	ta1s	• Name of the current tenant, which can contain digits, letters, and underscores (_). • Plan a tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or YARN queue that exists in the current cluster.
Tenant Type	Non-leaf Tenant	Whether the tenant is a leaf tenant. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Dynamic Resource	Yarn	Dynamic compute resources for the current tenant. • When **Yarn** is selected, the system automatically creates a queue in YARN and assigns the queue the same name as the tenant. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.
Default Resource Pool Capacity (%)	90	Percentage of resources used by the current tenant. The base value is the total resources of the parent tenant.
Default Resource Pool Max. Capacity (%)	80	Maximum percentage of compute resources used by the current tenant. The base value is the total resources of the parent tenant.
Storage Resource	HDFS	Storage resources selected for the current tenant. • When **HDFS** is selected, the system automatically creates a folder named after the sub-tenant in the HDFS parent tenant directory. • If **HDFS** is not selected, the system does not create a storage directory under the root directory of HDFS. If the parent tenant does not have storage resources, the sub-tenant cannot use storage resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Space Quota (MB)	500 MB	Quota for HDFS storage space used by the current tenant. ● The minimum value is **1**, and the maximum value is the total storage quota of the parent tenant. The unit is MB or GB. ● This parameter indicates the maximum HDFS storage space that can be used by a tenant, but does not indicate the actual space used. ● If its value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. ● If this quota is greater than the quota of the parent tenant, the actual storage space does not exceed the quota of the parent tenant. To ensure data reliability, the system automatically generates one backup file when a file is stored in the HDFS. That is, two replicas of the same file are stored by default. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value is set to **500** MB, the actual space for storing files is about 250 MB (500/2 = 250).
Storage Path	/tenant/ta1/ta1s	HDFS storage directory for the tenant. ● The system automatically creates a folder named after the sub-tenant name in the directory of the parent tenant by default. For example, if the sub-tenant is **ta1s** and the parent directory is **tenant/ta1**, the system sets this parameter for the sub-tenant to **tenant/ta1/ta1s**. ● The storage path is customizable in the parent directory. The parent directory for the storage path must be the storage directory of the parent tenant.
Service	-	Click **Associate Service** and select a service name from the **Service** drop-down list box. If **Association Mode** is set to **Exclusive**, service resources are occupied exclusively. If **share** is selected, service resources are shared.
Description	-	Description of the tenant.

Step 3 Click **OK** to save the settings.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

It takes a few minutes to save the settings. If the **Tenant created successfully** is displayed in the upper-right corner, the tenant is added successfully. The tenant is created successfully.

- Roles, compute resources, and storage resources are automatically created when tenants are created.
- The new role has permissions on the compute and storage resources. This role and its permissions are automatically controlled by the system and cannot be manually managed by choosing **System** > **Permission** > **Role**. The role name is in the format of *Tenant name_Cluster ID*. The ID of the first cluster is not displayed by default.
- When using this tenant, create a system user and assign the tenant role to the user. For details, see **Binding a Tenant to an MRS Cluster User**.
- The sub-tenant can further allocate the resources of its parent tenant. The sum of the resource percentages of direct sub-tenants under a parent tenant at each level cannot exceed 100%. The sum of the compute resource percentages of all level-1 tenants cannot exceed 100%.

----End

8.9.3.3 Binding a Tenant to an MRS Cluster User

Scenarios

A newly created tenant cannot directly log in to the cluster to access resources. You need to add a user for the tenant on FusionInsight Manager and bind the user to the role of the tenant to assign operation permissions to the user.

This section uses MRS clusters version 3.x or later as an example.

Prerequisites

You have clarified service requirements and created a tenant.

Binding a Tenant to an MRS Cluster User on the Console

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 If you want to add a user to the system, click **Create**.

If you want to bind tenant roles to an existing user in the system, locate the row of the user and click **Modify** in the **Operation** column.

Set user attributes according to **Table 8-54**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-54 User parameters

Parameter	Description
Username	Indicates the current username. The value contains 3 to 32 characters, including digits, letters, underscores (_), hyphens (-), and spaces. • The username cannot be the same as the OS username of any node in the cluster. Otherwise, the user cannot be used. • A username that differs only in alphabetic case from an existing username is not allowed. For example, if **User1** has been created, you cannot create **user1**. Enter the correct username when using **User1**.
User Type	The options are **Human-Machine** and **Machine-Machine**. • **Human-Machine** user: used for FusionInsight Manager O&M and component client operations. If you select this option, set both **Password** and **Confirm Password** accordingly. • **Machine-Machine** user: used for application development. If you select this option, the password is randomly generated.
Password	This parameter is mandatory if **User Type** is set to **Human-Machine**. The password must contain 8 to 64 characters of at least four types of the following: uppercase letters, lowercase letters, digits, special characters, and spaces. The password cannot be the username or the username spelled backwards.
Confirm Password	Enter the password again.
User Group	In the **User Group** area, click **Add** and select user groups to add the user to the groups. • If roles have been added to the user groups, the user can be granted the permissions of the roles. • For example, add the user to the Hive user group to assign Hive permissions to the user.
Primary Group	Select a group as the primary group for the user to create directories and files. The drop-down list contains all groups selected in **User Group**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Role	Click **Add** to bind a tenant role to the user. • To use the resources of tenant **tenant1** and add or delete its sub-tenants, a user must be assigned both the **Manager_tenant** role and the **tenant1**_*Cluster ID* role. • If the tenant is associated with the HBase service and Ranger authentication is enabled for the cluster, you must configure HBase execution permissions on the Ranger page.
Description	Indicates the description of the current user.

Step 4 Click **OK**.

----End

8.9.3.4 Adding an MRS Tenant Resource Pool

In an MRS cluster, users can logically divide Yarn cluster nodes to combine multiple NodeManagers into a Yarn resource pool. Each NodeManager belongs to one resource pool only. The system contains a **default** resource pool by default. All NodeManagers that are not added to customized resource pools belong to this resource pool.

You can create a customized resource pool on MRS and add hosts that have not been added to other customized resource pools to it.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Adding a Resource Pool on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Tenants**.

Step 4 Click the **Resource Pools** tab.

Step 5 Click **Create Resource Pool**.

Step 6 In **Create Resource Pool**, set the properties of the resource pool.

- **Name**: Enter a name for the resource pool. The name of the newly created resource pool cannot be **default**. The name can contain digits, letters, and underscores (_), and cannot start with an underscore (_).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Resource Label**: Enter the resource label of the resource pool. The value can contain 1 to 50 characters, including digits, letters, underscores (_), and hyphens (-), and must start with a digit or letter.
- **Available Hosts**: In the host list on the left, select a specified host name and add it to the resource pool. Only hosts in the cluster can be selected. The host list of a resource pool can be left blank.

Step 7 Click **OK**.

Step 8 After a resource pool is created, users can view the **Name**, **Members**, **Type**, **vCore** and **Memory** in the resource pool list. Hosts that are added to the customized resource pool are no longer members of the **default** resource pool.

----End

Adding a Resource Pool on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Tenant Resources**.

For MRS 2.x or earlier versions, select **Tenant**.

Step 3 Click **Resource Pool** in the navigation pane, and click **Add Resource Pool**.

Step 4 Set resource pool properties.

- **MRS cluster 3.*x* and later versions:**
  - **Cluster**: Select the cluster to which the resource pool is to be added. (This parameter is unavailable for clusters of MRS 3.3.0 or later.)
  - **Name**: Enter a name for the resource pool. The value can contain digits, letters, and underscores (_), and cannot start with an underscore (_).
  - **Resource Label**: Enter the resource label of the resource pool. The value can contain 1 to 50 characters, including digits, letters, underscores (_), and hyphens (-), and must start with a digit or letter.
  - **Resource**: In the **Available Hosts** area, select specified hosts and click **>>** to add the hosts to the **Selected Hosts** area. Only hosts in the cluster can be selected. The host list of a resource pool can be left blank.

    You can search for hosts by host name, number of CPU cores, memory, operating system, or platform type based on service requirements.
- **MRS 2.x and earlier versions:**
  - **Name**: Enter a name for the resource pool. The name of the newly created resource pool cannot be **Default**.

    The name consists of 1 to 20 characters and can contain digits, letters, and underscores (_) but cannot start with an underscore (_).
  - **Hosts**: In the host list on the left, select the name of a specified host and click **>>** to add the selected host to the resource pool. Only hosts in the cluster can be selected. The host list of a resource pool can be left blank.

Step 5 Click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the resource pool is created, you can view it including the name, members, and mode in the resource pool list. Hosts that are added to the customized resource pool are no longer members of the **default** resource pool.

----End

8.9.3.5 Configuring the Queue Capacity Policy of a Resource Pool

After a resource pool is added, the capacity policies of available resources need to be configured for Yarn task queues. This ensures that tasks in the resource pool are running properly. Each queue can be configured with the queue capacity policy of only one resource pool. Users can view the queues in any resource pool and configure queue capacity policies. After the queue policies are configured, Yarn task queues and resource pools are associated.

You can configure queue policies on MRS.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.
- A resource pool has been added.
- The task queue is not associated with other resource pools except the default **default** resource pool.

Configuring a Queue Capacity Policy on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Tenants**.

Step 4 Click the **Resource Distribution Policies** tab.

Step 5 In **Resource Pools**, select a specified resource pool.

Available Resource Quota: indicates that all resources in each resource pool are available for queues by default.

Step 6 Locate the row that contains the target queue in the **Resource Allocation** area, and click **Modify** in the **Operation** column.

Step 7 In **Modify Resource Allocation**, configure the resource capacity policy of the task queue in the resource pool.

- **Capacity (%)**: specifies the percentage of the current tenant's computing resource usage.
- **Maximum Capacity (%)**: specifies the percentage of the current tenant's maximum computing resource usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click **OK** to save the settings.

----End

Configuring the Queue Capacity Policy on Manager

For clusters of MRS 3.*x* and later

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Tenant Resources** > **Dynamic Resource Plan** and click the **Resource Distribution Policy** tab.

Step 3 Select the name of the target cluster from **Cluster** and select a resource pool from **Resource Pool**.

Step 4 Locate the row that contains the target queue in the **Resource Allocation** area, and click **Modify** in the **Operation** column.

Step 5 Modify the resource allocation.

- **Capacity scheduler**
  a. In the **Modify Resource Allocation** window, click **Resource Configuration Policy** and set the resource capacity policy for the task queue in the resource pool.
     - **Capacity (%)**: indicates the percentage of compute resources used by the current tenant.
     - **Maximum Capacity (%)**: indicates the maximum percentage of compute resources used by the current tenant.
- **Superior Scheduler**
  a. In the **Modify Resource Allocation** window, click **Resource Configuration Policy** and configure a resource configuration policy for the queue in the resource pool.
     - **Weight**: The task queue with a larger weight preempts resources first when resources are insufficient. Its initial value is the same as the minimum resource percentage.
     - **Minimum Resource**: indicates the minimum resources that a tenant can obtain.
     - **Maximum Resource**: indicates the maximum resources that a tenant can obtain.
     - **Reserved Resource**: indicates the resources that are reserved for the tenant's queues and cannot be lent to other tenants' queues.
  b. In the **Modify Resource Allocation** dialog box, click the **User Policy** tab and set the user policy.

     defaultUser(built-in) indicates that the policy specified by **defaultUser** is applied if no policy is explicitly specified by the user. The default policy cannot be deleted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Click **Add User Policy** to add a user policy.
  - **Username**: indicates the name of a user.
  - **Weight**: The task queue with a larger weight preempts resources first when resources are insufficient. (available for versions earlier than MRS 3.5.0)
  - **Priority**: When task queues need to compete for resources, a queue with a higher priority obtains resources first. However, not all containers can be started. (available for MRS 3.5.0 and later)
  - **Max vCores**: indicates the maximum number of virtual cores that the user can obtain.
  - **Max Memory(MB)**: indicates the maximum memory that the user can obtain.
- Click **Modify** in the **Operation** column to modify an existing user policy.
- Click **Clear** in the **Operation** column to delete an existing user policy.

Step 6 Click **OK** to save the settings.

----End

For MRS clusters of 2.x and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Tenant**.

Step 3 Click the **Dynamic Resource Plan** tab.

Step 4 In **Resource Pools**, select a specified resource pool.

Available Resource Quota: indicates that all resources in each resource pool are available for queues by default.

Step 5 Locate the row that contains the target queue in the **Resource Allocation** area, and click **Modify** in the **Operation** column.

Step 6 In **Modify Resource Allocation**, configure the resource capacity policy of the task queue in the resource pool.

- **Capacity (%)**: specifies the percentage of the current tenant's computing resource usage.
- **Maximum Capacity (%)**: specifies the percentage of the current tenant's maximum computing resource usage.

Step 7 Click **OK** to save the settings.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.9.3.6 Configuring the MRS Tenant Queue

You can modify the queue configuration of a specified tenant on MRS to scale in or out the queue based on service requirements. YARN queues are associated with resource pools for resource allocation and scheduling.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.
- A tenant associated with Yarn and allocated dynamic resources has been added.

Configuring a Tenant Queue on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the MRS details page, click **Tenants**.

Step 4 Click the **Queue Configuration** tab.

Step 5 In the tenant queue table, click **Modify** in the **Operation** column of the specified tenant queue.

In the tenant list on the left of the **Tenants** tab, click the target tenant. In the window that is displayed, choose **Resources**. On the page that is displayed, click  to open the queue modification page.

A queue can be bound to only one non-default resource pool.

- **For MRS 2.x and earlier versions:**

Table 8-55 Queue configuration parameters

Parameter	Example Value	Description
Maximum Applications	10	Specifies the maximum number of applications.
Maximum AM Resource Percent	10	Specifies the maximum percentage of resources that can be used to run the ApplicationMaster in a cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Minimum User Limit Percent (%)	25	Specifies the minimum percentage of resources consumed by a user. The resources for each user in a queue are limited at any time. If applications of multiple users are running at the same time in a queue, the resource usage of each user fluctuates between the minimum value and the maximum value. The minimum value is determined by the number of running applications, while the maximum value is determined by this parameter. For example, assume that this parameter is set to **25**. If two users submit applications to the queue, each user can use a maximum of 50% resources; if three users submit applications to the queue, each user can use a maximum of 33% resources; if four users submit applications to the queue, each user can use a maximum of 25% resources.
User Limit Factor	10	Specifies the limit factor of the maximum user resource usage. The maximum user resource usage percentage can be obtained by multiplying the limit factor with the percentage of the tenant's actual resource usage in the cluster.
Status	-	Specifies the current status of a resource plan.
Default Resource Pool (Default Node Label Expression)	default	Specifies the resource pool used by a queue. The default value is **default**. If you want to change the resource pool, configure the queue capacity first. For details, see **Configuring the Queue Capacity Policy of a Resource Pool**.

- **For MRS cluster 3.*x* and later versions:**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-56 Queue configuration parameters

Parameter	Example Value	Description
Max Master Shares (%)	10	Indicates the maximum percentage of resources occupied by all ApplicationMasters in the current queue.
Max Allocated vCores	-1	Indicates the maximum number of cores that can be allocated to a single YARN container in the current queue. The default value is **-1**, indicating that the number of cores is not limited within the value range.
Max Allocated Memory (MB)	-1	Indicates the maximum memory that can be allocated to a single Yarn container in the current queue. The default value is **-1**, indicating that the memory is not limited within the value range.
Max Running Apps	-1	Indicates the maximum number of tasks that can be executed at the same time in the current queue. The default value is **-1**, indicating that the number is not limited within the value range (the meaning is the same if the value is empty). The value 0 indicates that the task cannot be executed.
Max Running Apps per User	-1	Maximum number of tasks that can be executed by each user in the current queue at the same time. The default value is **-1**, indicating that the number is not limited within the value range. If the value is **0**, the task cannot be executed.
Max Pending Apps	-1	Indicates the maximum number of tasks that can be suspended at the same time in the current queue. The default value is **-1**, indicating that the number is not limited within the value range (the meaning is the same if the value is empty). The value **0** indicates that tasks cannot be suspended. The maximum number of pending tasks cannot be less than the number of jobs submitted to the current queue in the scheduling period (100 ms).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Resource Allocation Rule	FIFO	Indicates the rule for allocating resources to different tasks of a user. The rule can be FIFO or FAIR. If a user submits multiple tasks in the current queue and the rule is FIFO, the tasks are executed one by one in sequential order. If the rule is FAIR, resources are evenly allocated to all tasks.
Default Resource Label	-	Indicates that tasks are executed on a node with a specified resource label. If you need to use a new resource pool, change the default label to the new resource pool label.
Cross-Pool Scheduling	-	Indicates whether containers in the current queue support cross-pool scheduling. This parameter is available for clusters of MRS 3.3.0 to 3.5.0. This function cannot be enabled for the default queue.
Cross-Pool AM Scheduling	-	Indicates whether ApplicationMasters in the current queue support cross-pool scheduling. This parameter is available for clusters of MRS 3.3.0 to 3.5.0. This function cannot be enabled for the default queue.
Active	ACTIVE	– **ACTIVE**: indicates that the current queue can receive and execute tasks. – **INACTIVE**: indicates that the current queue can receive but cannot execute tasks. Tasks submitted to the queue are suspended.
Open	OPEN	– **OPEN**: indicates that the current queue is opened. – **CLOSED**: indicates that the current queue is closed. Tasks submitted to the queue are rejected.

----End

Configuring a Tenant Queue on Manager

For clusters of MRS 3.*x* and later

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click *Tenant Resources*.

Step 3 Click the **Dynamic Resource Plan** tab.

Step 4 Click the **Queue Configuration** tab.

Step 5 Locate the row containing the specified tenant resource name and click **Modify** in the **Operation** column. Modify the parameters based on the scheduler type in use.

NOTE

- YARN in a new cluster uses the Superior scheduler by default. You can switch the scheduler by referring to **Switching the MRS Tenant Resource Scheduler**.
- To query the scheduler type, log in to Manager and search for the **yarn.resourcemanager.scheduler.class** parameter on the **All Configurations** page of the YARN service.
- You can also access the **Modify Queue Configuration** page as follows: In the tenant list on the **Tenant Resources Management** page, click the target tenant, click the **Resource** tab, and click next to **Queue Configurations (*Queue name*)**.
- A queue can be bound to only one non-default resource pool. That is, a newly added resource pool can be bound to only one queue to serve as the default resource pool of the queue.
- For parameters such as **Max Allocated vCores**, **Max Allocated Memory(MB)**, **Max Running Apps**, **Max Running Apps per User**, and **Max Pending Apps**, if the value of a sub-tenant is **-1**, the value of the parent tenant can be set to a specific limit. If the parent tenant value is a specific limit, the sub-tenant value can be set to **-1**.
- **Max Allocated vCores** and **Max Allocated Memory(MB)** must be both changed to values other than **-1**.
- For queues with cross-resource-pool scheduling enabled, existing resource pools cannot be deleted during job running. Otherwise, running jobs may be continuously blocked because they cannot obtain resources. Similarly, if a new resource pool is configured for a queue during job running, the queue in the running state may not immediately use the resources in the new resource pool. The new resources are available only for jobs submitted after modification.
- To use a Superior scheduler cluster, configure sub-tenant properties by referring to the following table.

Table 8-57 Queue configuration parameters

Parameter	Example Value	Description
Max Master Shares (%)	10	Indicates the maximum percentage of resources occupied by all ApplicationMasters in the current queue.

Parameter	Example Value	Description
Max Allocated vCores	-1	Indicates the maximum number of cores that can be allocated to a single YARN container in the current queue. The default value is **-1**, indicating that the number of cores is not limited within the value range.
Max Allocated Memory (MB)	-1	Indicates the maximum memory that can be allocated to a single Yarn container in the current queue. The default value is **-1**, indicating that the memory is not limited within the value range.
Max Running Apps	-1	Indicates the maximum number of tasks that can be executed at the same time in the current queue. The default value is **-1**, indicating that the number is not limited within the value range (the meaning is the same if the value is empty). The value 0 indicates that the task cannot be executed.
Max Running Apps per User	-1	Maximum number of tasks that can be executed by each user in the current queue at the same time. The default value is **-1**, indicating that the number is not limited within the value range. If the value is **0**, the task cannot be executed.
Max Pending Apps	-1	Indicates the maximum number of tasks that can be suspended at the same time in the current queue. The default value is **-1**, indicating that the number is not limited within the value range (the meaning is the same if the value is empty). The value **0** indicates that tasks cannot be suspended. The maximum number of pending tasks cannot be less than the number of jobs submitted to the current queue in the scheduling period (100 ms).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Resource Allocation Rule	FIFO	Indicates the rule for allocating resources to different tasks of a user. The rule can be FIFO or FAIR. If a user submits multiple tasks in the current queue and the rule is FIFO, the tasks are executed one by one in sequential order. If the rule is FAIR, resources are evenly allocated to all tasks.
Default Resource Label	-	Indicates that tasks are executed on a node with a specified resource label. If you need to use a new resource pool, change the default label to the new resource pool label.
Cross-Pool Scheduling	-	Indicates whether containers in the current queue support cross-pool scheduling. This parameter is available for clusters of MRS 3.3.0 to 3.5.0. This function cannot be enabled for the default queue.
Cross-Pool AM Scheduling	-	Indicates whether ApplicationMasters in the current queue support cross-pool scheduling. This parameter is available for clusters of MRS 3.3.0 to 3.5.0. This function cannot be enabled for the default queue.
Active	ACTIVE	– **ACTIVE**: indicates that the current queue can receive and execute tasks. – **INACTIVE**: indicates that the current queue can receive but cannot execute tasks. Tasks submitted to the queue are suspended.
Open	OPEN	– **OPEN**: indicates that the current queue is opened. – **CLOSED**: indicates that the current queue is closed. Tasks submitted to the queue are rejected.

- For a cluster using the Capacity scheduler, **Tenant Resource Name (Queue)** indicates the tenant and queue name. Configure sub-tenants properties by referring to the following table.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 8-58 Queue configuration parameters

Parameter	Example Value	Description
Maximum Applications	10	Specifies the maximum number of applications.
Maximum AM Resource Percent	10	Specifies the maximum percentage of resources that can be used to run the ApplicationMaster in a cluster.
Minimum User Limit Percent (%)	25	Specifies the minimum percentage of resources consumed by a user. The resources for each user in a queue are limited at any time. If applications of multiple users are running at the same time in a queue, the resource usage of each user fluctuates between the minimum value and the maximum value. The minimum value is determined by the number of running applications, while the maximum value is determined by this parameter. For example, assume that this parameter is set to **25**. If two users submit applications to the queue, each user can use a maximum of 50% resources; if three users submit applications to the queue, each user can use a maximum of 33% resources; if four users submit applications to the queue, each user can use a maximum of 25% resources.
User Limit Factor	10	Specifies the limit factor of the maximum user resource usage. The maximum user resource usage percentage can be obtained by multiplying the limit factor with the percentage of the tenant's actual resource usage in the cluster.
Status	-	Specifies the current status of a resource plan.
Default Resource Pool (Default Node Label Expression)	default	Specifies the resource pool used by a queue. The default value is **default**. If you want to change the resource pool, configure the queue capacity first. For details, see **Configuring the Queue Capacity Policy of a Resource Pool**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Click **OK**.

----End

For MRS clusters of 2.x and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Tenant**.

Step 3 Click the **Dynamic Resource Plan** tab.

Step 4 Click the **Queue Configuration** tab.

Step 5 In the tenant queue table, click **Modify** in the **Operation** column of the specified tenant queue.

In the tenant list on the left of the **Tenant** tab, click the target tenant. In the window that is displayed, choose **Resource**. On the page that is displayed, click the edit icon to open the queue modification page.

Table 8-59 Queue configuration parameters

Parameter	Example Value	Description
Maximum Applications	10	Specifies the maximum number of applications.
Maximum AM Resource Percent	10	Specifies the maximum percentage of resources that can be used to run the ApplicationMaster in a cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Minimum User Limit Percent (%)	25	Specifies the minimum percentage of resources consumed by a user. The resources for each user in a queue are limited at any time. If applications of multiple users are running at the same time in a queue, the resource usage of each user fluctuates between the minimum value and the maximum value. The minimum value is determined by the number of running applications, while the maximum value is determined by this parameter. For example, assume that this parameter is set to **25**. If two users submit applications to the queue, each user can use a maximum of 50% resources; if three users submit applications to the queue, each user can use a maximum of 33% resources; if four users submit applications to the queue, each user can use a maximum of 25% resources.
User Limit Factor	10	Specifies the limit factor of the maximum user resource usage. The maximum user resource usage percentage can be obtained by multiplying the limit factor with the percentage of the tenant's actual resource usage in the cluster.
Status	-	Specifies the current status of a resource plan.
Default Resource Pool (Default Node Label Expression)	default	Specifies the resource pool used by a queue. The default value is **default**. If you want to change the resource pool, configure the queue capacity first. For details, see **Configuring the Queue Capacity Policy of a Resource Pool**.

----End

8.9.4 Managing MRS Tenant Resources

8.9.4.1 Managing the MRS Tenant Resource Directory

You can manage the HDFS storage directory used by a specific tenant on MRS. The management operations include adding a tenant directory, modifying the directory file quota, modifying the storage space, and deleting a directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.
- A tenant associated with HDFS storage resources has been added.

Managing a Tenant Directory

Step 1 Go to the **Tenant** page.

- On the MRS management console: Click the target cluster. On the cluster details page, click **Tenant**.
- On the MRS Manager of an MRS 3.*x* or later cluster: Click **Tenant Resources**.
- On the MRS Manager of an MRS 2.*x* or earlier cluster: Click **Tenant**.

Step 2 In the tenant list on the left, click the target tenant.

Step 3 Click the **Resource** tab and perform the following operations to manage tenant directories:

- **View the tenant directory.**

  View related information in the **HDFS Storage** pane.

  – The **File Number Threshold** column provides the quota for files and directories of the tenant directory. (On the MRS management console, this parameter is **Maximum Number of Files/Directories**.)

  – The **Storage Space Quota** column indicates storage space size of tenant directories.

- **Add a tenant directory.**

  a. In the **HDFS Storage** pane, click **Create Directory**.

  - **Parent Directory**: Select a storage directory of a parent tenant.

    This parameter is not displayed if the current tenant is not a sub-tenant. If the parent tenant has multiple directories, select any of them.

  - **Path**: Enter the tenant directory.
    - If the current tenant is not a sub-tenant, the new path is created in the HDFS root directory.
    - If the current tenant is a sub-tenant, the new path is created in the specified directory.

    A complete HDFS storage directory can contain a maximum of 1,023 characters. An HDFS directory name contains digits, letters, spaces, and underscores (_). The name cannot start or end with a space.

  - **Maximum Number of Files/Directories**: Enter the quota of files and directories. This parameter is optional.

  - **File Number Threshold (%)** is valid only when **Quota** is set. If the ratio of the number of used files to the value of **Quota** exceeds the value of this parameter, an alarm is generated. If this parameter is

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

not specified, no alarm is reported in this scenario. This parameter is available for clusters of MRS 3.x or later.

The number of used files is collected hourly. Consequently, the alarm indicating that the usage ratio exceeds the threshold is triggered with a delay.

- **Storage Space Quota**: Specify the storage space size of the tenant directory.

  To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two replicas are generated in total. The HDFS storage space indicates the total disk space occupied by all these replicas. For example, if the value of **Storage Space Quota** is set to **500**, the actual space for storing files is about 250 MB (500/2 = 250).

- Storage Space Threshold (%): If the ratio of used storage space to the value of **Space Quota** exceeds this value, an alarm is generated. If this parameter is not specified, no alarm is reported in this scenario. This parameter is available for clusters of MRS 3.x or later.

  The used storage space is collected hourly. Consequently, the alarm indicating that the usage ratio exceeds the threshold is triggered with a delay.

b. Click **OK**. The system creates tenant directories in the HDFS root directory.

- **Modify a tenant directory.**

  a. In the **HDFS Storage** table, click **Modify** in the **Operation** column of the specified tenant directory.

  b. Change the parameter values based on service requirements.

  c. Click **OK**.

- **Delete a tenant directory.**

  a. In the **HDFS Storage** table, click **Delete** in the **Operation** column of the specified tenant directory.

  The default HDFS storage directory set during tenant creation cannot be deleted. Only the newly added HDFS storage directory can be deleted.

  b. Click **OK**. The tenant directory is deleted.

----End

8.9.4.2 Managing MRS Tenant Resource Pools

When hosts in a resource pool need to be adjusted based on service requirements, you can modify members in the resource pool on MRS.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- To delete a resource pool, the following requirements must be met:
  - Any queue in a cluster cannot use the resource pool to be deleted as the default resource pool. Before deleting the resource pool, cancel the default resource pool. For details, see **Configuring the MRS Tenant Queue**.
  - Resource distribution policies of all queues have been cleared from the resource pool being deleted. For details, see **Clearing the MRS Tenant Queue Configuration**.

Managing Resource Pools on the MRS Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click **Tenants**.

Step 4 Click the **Resource Pools** tab.

Step 5 Locate the row that contains the specified resource pool, and click **Modify** in the **Operation** column.

Step 6 Manage hosts in a resource pool.

- Adding a host: In the host list on the left, select the specified host name and add it to the resource pool.
- Deleting a host: In the host list on the right, click  next to a host to remove the host from the resource pool. The host list of a resource pool can contain no hosts.

Step 7 Click **OK**.

Locate the row that contains the target resource pool in the **Resource Pools** tab, and click **Delete** in the **Operation** column. In the displayed dialog box, click **OK**.

Step 8 Check whether there are running resident jobs, such as Flink streaming jobs, on the target host.

- If yes, perform the following steps to decommission and recommission the NodeManager instances on the hosts whose resource pool has changed:
  a. On MRS Manager, choose **Cluster** > **Services** > **Yarn** > **Instances**.
  b. Select the NodeManager instances that have the same host names as the hosts whose resource pool has changed.
  c. Choose **More** > **Decommission**, enter the user password, confirm the decommissioning impact, and click **OK**. Wait until the decommissioning completes successfully.
  d. Select the decommissioned instances, choose **More** > **Recommission**, enter the user password, confirm the recommissioning impact, and click **OK**. Wait until the recommissioning completes successfully.
- If no, no further action is required.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Managing Resource Pools on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Tenant Resources** > **Resource Pool**.

For MRS 2.x or earlier versions, choose **Tenant** > **Resource Pool**.

Step 3 Locate the row that contains the specified resource pool, and click **Edit** in the **Operation** column.

For MRS 2.x and earlier versions and MRS 3.3.0 and later versions, click **Modify**.

Step 4 Manage hosts in a resource pool.

- Adding a host: Select the name of a specified host in host list on the left and click **>>** to add the selected host to the resource pool.
- Deleting a host: In the host list on the right, select the name of a specified host and click **<<** to add the selected host to the resource pool. The host list of a resource pool can be left blank.

Step 5 Click **OK**.

Locate the row that contains the target resource pool in the **Resource Pools** tab, and click **Delete** in the **Operation** column. In the displayed dialog box, click **OK**.

Step 6 Check whether there are running resident jobs, such as Flink streaming jobs, on the target host.

- If yes, perform the following steps to decommission and recommission the NodeManager instances on the hosts whose resource pool has changed:
  a. On MRS Manager, choose **Cluster** > **Services** > **Yarn** > **Instances**.
  b. Select the NodeManager instances that have the same host names as the hosts whose resource pool has changed.
  c. Choose **More** > **Decommission**, enter the user password, confirm the decommissioning impact, and click **OK**. Wait until the decommissioning completes successfully.
  d. Select the decommissioned instances, choose **More** > **Recommission**, enter the user password, confirm the recommissioning impact, and click **OK**. Wait until the recommissioning completes successfully.
- If no, no further action is required.

----End

8.9.4.3 Clearing the MRS Tenant Queue Configuration

Users can clear the configuration of a queue on MRS Manager when the queue does not need resources from a resource pool or if a resource pool needs to be disassociated from the queue. Clearing queue configurations means that the resource capacity policy of the queue is canceled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.
- If a queue is to be unbound from a resource pool, this resource pool cannot serve as the default resource pool of the queue. Therefore, you must first change the default resource pool of the queue to another one. For details, see **Configuring the MRS Tenant Queue**.

Clearing Queue Configuration

Step 1 Access the resource distribution policy page.

- Log in to the MRS console.
- On the MRS console, select the target resource pool from the **Resource Pools** drop-down list.
- On the Manager, the page path is as follows:
  - For MRS 3.x and later versions, choose **Tenant Resources** > **Dynamic Resource Plan** > **Resource Distribution Policy**.
  - For MRS 2.x and earlier versions, choose **Tenant** > **Dynamic Resource Plan** > **Resource Distribution Policy**.

Step 2 Locate the specified queue in the **Resource Allocation** table, and click **Clear** in the **Operation** column.

In the displayed dialog box, click **Yes** to clear the queue configurations from the current resource pool.

If no resource capacity policy is configured for a queue, the clearing function is unavailable for that queue by default.

----End

8.9.4.4 Restoring MRS Tenant Data After YARN Is Reinstalled

Tenant data is stored on Manager and in cluster components by default. When components are restored from faults or reinstalled, some tenant configuration data may be abnormal. In this case, you can manually restore the tenant data.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Restoring Tenant Data

Step 1 Go to the **Tenant** page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- On the MRS management console: Click the target cluster. On the cluster details page, click **Tenant**.
- On the MRS Manager of an MRS 3.*x* or later cluster: Click **Tenant Resources**.
- On the MRS Manager of an MRS 2.*x* or earlier cluster: Click **Tenant**.

Step 2 In the tenant list on the left, click a tenant node.

Step 3 Check the status of the tenant data.

1. In the **Summary** tab, check **Tenant Status**. A green icon indicates that the tenant is available and gray indicates that the tenant is unavailable.
2. Click **Resources** and check the status of **Yarn** or **HDFS Storage**. Green indicates that the resource is available, and gray indicates that the resource is unavailable.
3. Click **Service Association** and check the **Status** column of the associated service table. **Good** indicates that the component can provide services for the associated tenant. **Bad** indicates that the component cannot provide services for the tenant.
4. If any check result is abnormal, go to **Step 4** to restore tenant data.

Step 4 Click **Restore Tenant Data**.

To restore data on the Manager of an MRS 3.x or later cluster, click . In the displayed dialog box, enter the password of the current user and click **OK**.

Step 5 In the displayed window, select one or more components whose data needs to be restored. Click **OK**. The system automatically restores the tenant data.

----End

8.9.4.5 Deleting an MRS Tenant

You can delete MRS tenants that are no longer used based on service requirements, releasing resources occupied by the tenants.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.
- You have checked whether the tenant to be deleted has sub-tenants. If the tenant has sub-tenants, delete them; otherwise, you cannot delete the tenant.
- The role of the tenant to be deleted cannot be associated with any user or user group.

Deleting a Tenant on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 On the cluster details page, click **Tenants**.

Step 3 In the tenant list on the left, move the cursor to the tenant node to be deleted and click **Delete**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The **Delete Tenant** dialog box is displayed. If you want to save the tenant data, select **Reserve the data of this tenant**. Otherwise, the tenant's storage space will be deleted.

Step 4 Click **OK**.

It takes a few minutes to save the configuration. After the tenant is deleted successfully, the tenant's role and storage space are also removed.

- Deleting a tenant does not automatically remove its task queue from YARN.
- If you choose not to reserve data when deleting the parent tenant, the data of sub- tenants that use storage resources is also deleted.

----End

Deleting a Tenant on Manager

For MRS 3.*x* and later versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Tenant Resources**.

Step 3 In the tenant list on the left, click the target tenant and click .

- If you want to save the tenant data, select **Reserve the data of this tenant**. Otherwise, the tenant's storage space will be deleted.
- To delete a tenant without retaining its data as a user outside the supergroup, first log in to the HDFS client as a supergroup user, and then manually clear the tenant's storage space to avoid residual data.

Step 4 In the **Delete Tenant** dialog box, enter **DELETE** in the confirmation text box. Click **OK**.

It takes a few minutes to save the configuration. After the tenant is deleted successfully, the tenant's role and storage space are also removed.

Deleting a tenant does not automatically remove its task queue from YARN. The tenant's queue is not displayed on the role management page in YARN.

----End

For MRS 2.x and earlier versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Tenant**.

Step 3 In the tenant list on the left, move the cursor to the tenant node to be deleted and click **Delete**.

The **Delete Tenant** dialog box is displayed. If you want to save the tenant data, select **Reserve the data of this tenant**. Otherwise, the tenant's storage space will be deleted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Click **OK**.

It takes a few minutes to save the configuration. After the tenant is deleted successfully, the tenant's role and storage space are also removed.

- Deleting a tenant does not automatically remove its task queue from YARN.
- If you choose not to reserve data when deleting the parent tenant, the data of sub- tenants that use storage resources is also deleted.

----End

8.9.4.6 Managing Global User Policies When Using Superior Scheduler

Scenarios

If a tenant uses a Superior scheduler, you can configure the global policy for users to use the resource scheduler, including:

- Maximum running apps
- Maximum pending apps
- Default queue

Managing MRS Global User Policies

- Add a policy.
  a. Log in to FusionInsight Manager of the MRS cluster.

     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
  b. Click **Tenant Resources**.
  c. Choose **Dynamic Resource Plan**.
  d. Click the **Global User Policy** tab.

     defaults(default setting) indicates that the policy specified for **defaults** is applied if a user does not have a global policy. The default policy cannot be deleted.
  e. Click **Create Global User Policy**. In the displayed dialog box, set the following parameters:

     - **Cluster**: Select the target cluster. (This parameter is unavailable for clusters of MRS 3.3.0 or later.)
     - **Username**: indicates the user for whom resource scheduling is controlled. Enter an existing username in the current cluster.
     - **Max Running Apps**: indicates the maximum number of tasks that the user can run in the current cluster.
     - **Max Pending Apps**: indicates the maximum number of tasks that the user can suspend in the current cluster.
     - **Default Queue**: indicates the queue of the user. Enter the name of an existing queue in the current cluster.
- Modify a policy.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

a. Log in to FusionInsight Manager of the MRS cluster.

   For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

b. Click **Tenant Resources**.

c. Choose **Dynamic Resource Plan**.

d. Click the **Global User Policy** tab.

e. In the row that contains the desired user policy, click **Modify** in the **Operation** column.

f. In the displayed dialog box, modify parameters and click **OK**.

- Delete a policy.

  a. Log in to FusionInsight Manager of the MRS cluster.

     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

  b. Click **Tenant Resources**.

  c. Choose **Dynamic Resource Plan**.

  d. Click the **Global User Policy** tab.

  e. In the row that contains the desired user policy, click **Delete** in the **Operation** column.

     In the displayed dialog box, click **OK**.

8.9.4.7 Clearing Tenant's Non-Associated Queues Using Capacity Scheduler

Scenarios

If Yarn uses the Capacity scheduler, deleting a tenant only sets the queue capacity of the tenant to **0** and the tenant status to **STOPPED** but does not clear the queues of the tenant in Yarn. Limited by the Yarn mechanism, queues cannot be dynamically deleted. You can run commands to manually delete residual queues.

Impact on the System

- During the script execution, the Controller service is restarted, Yarn configurations are synchronized, and the active and standby ResourceManagers are restarted.
- FusionInsight Manager becomes inaccessible during the restart of the Controller service.
- After the active and standby ResourceManagers are restarted, an alarm is generated indicating that Yarn and components that depend on Yarn are temporarily unavailable.

Prerequisites

Queues of a deleted tenant still exist.

Clearing Non-Associated Queues of a Tenant

Step 1 Check that queues of the deleted tenant still exist.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to MRS Manager and choose **Cluster** > **Services** > **Yarn**. Click the link of the active ResourceManager in **ResourceManager WebUI** to go to the ResourceManager web UI.

   For details about how to log in to MRS Manager, see **Accessing MRS Manager**.
2. Click **Scheduler** in the navigation tree on the left. In the right pane, you can view that queues of the tenant still exist in the **STOPPED** state and their **Configured Capacity** is **0**.

Step 2 Log in to the active management node as user **omm**.

Step 3 Switch the directory and execute the **cleanQueuesAndRestartRM.sh** script.

cd ${BIGDATA_HOME}/om-server/om/sbin

./cleanQueuesAndRestartRM.sh -c *Cluster ID*

You can choose **Cluster**, click the cluster name, and choose **Cluster Properties** on FusionInsight Manager to view the cluster ID.

During the script execution, you need to enter **yes** and the password.

```
Running the script will restart Controller and restart ResourceManager.
Are you sure you want to continue connecting (yes/no)?yes
Please input admin password:
Begin to backup queues ...
...
```

Step 4 After the script is executed successfully, log in to FusionInsight Manager, choose **Cluster**, click the cluster name, and choose **Services** > **Yarn**. Click the link of the active ResourceManager in **ResourceManager WebUI** to go to the ResourceManager web UI.

Step 5 Click **Scheduler** in the navigation tree on the left. In the right pane, you can view that queues of the tenant have been cleared.

----End

8.9.5 Switching the MRS Tenant Resource Scheduler

Scenarios

A newly installed MRS cluster uses the Superior scheduler by default. A cluster administrator can switch to other schedulers as required based on the instructions provided below.

Prerequisites

- The network connectivity of the cluster is proper and secure, and the YARN service status is normal.
- During scheduler switching, tenants cannot be added, deleted, or modified. In addition, services cannot be started or stopped.

Switching Between the Capacity Scheduler and Superior Scheduler (Available for Clusters of MRS 3.3.0 or Later)

This function is only available for clusters of MRS 3.3.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Constraints

- **This operation is available exclusively for newly provisioned clusters when the scheduler needs to be changed.**
- During the scheduler switchover, do not perform any operation on the cluster. Otherwise, the operation may fail due to database modification.

Impact on the system

- Because the ResourceManager is restarted during scheduler switching, submitting jobs to Yarn will fail at that time.
- After the scheduler is switched, the parameters of the scheduler that takes over the workload are used.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > **Yarn** and check whether the Yarn service status is normal. If the service is abnormal, restore the service.

Step 3 Log in to the active management node as user **omm**.

Step 4 Switch the scheduler.

- Run the following command to switch from the Capacity scheduler to the Superior scheduler:

  sh ${BIGDATA_HOME}/om-server/om/sbin/cleanSwitchScheduler.sh 1

  If information similar to the following is displayed, the switch is successful:

  ```
  Will change scheduler type to SUPERIOR
  Start to delete all tenant resource.
  End to delete all tenant resource.
  Start to delete all resource pool.
  End to delete all resource pool.
  ...
  End to switch scheduler by reset.
  ```

- Run the following command to switch from the Capacity scheduler to the Superior scheduler:

  sh ${BIGDATA_HOME}/om-server/om/sbin/cleanSwitchScheduler.sh 0

  If information similar to the following is displayed, the switch is successful:

  ```
  Will change scheduler type to CAPACITY
  Start to delete all tenant resource.
  End to delete all tenant resource.
  Start to delete all resource pool.
  End to delete all resource pool.
  ...
  End to switch scheduler by reset.
  ```

You can query the scheduler switching logs on the active management node.

- ${BIGDATA_LOG_HOME}/controller/aos/clean_switch_scheduler.log
- ${BIGDATA_LOG_HOME}/controller/aos/aos.log
- ${BIGDATA_LOG_HOME}/controller/aos/plugin.log

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Switching from the Capacity Scheduler to the Superior Scheduler

Impact on the system

- Because the ResourceManager is restarted during scheduler switching, submitting jobs to YARN will fail at that time.
- During scheduler switching, tasks in a job being executed on YARN will continue, but new tasks cannot be started.
- After scheduler switching is complete, jobs executed on YARN may fail, causing service interruptions.
- After scheduler switching is complete, parameters of the Superior scheduler are used for tenant management.
- After scheduler switching is complete, tenant queues whose capacity is 0 in the Capacity scheduler cannot be allocated resources in the Superior scheduler. As a result, jobs submitted to these tenant queues fail to be executed. Therefore, you are advised not to set the capacity of a tenant queue to 0 in the Capacity scheduler.
- After scheduler switching is complete, you cannot add or delete resource pools, YARN node labels, or tenants during the observation period. If such an operation is performed, the scheduler cannot be rolled back to the Capacity scheduler.

  The recommended observation period for scheduler switching is one week. If resource pools, YARN node labels, or tenants are added or deleted during this period, the observation period ends immediately.
- Rollback may cause the loss of partial or all Yarn job information.

Procedure

Step 1 Modify YARN service parameters and ensure that the YARN service status is normal.

1. Log in to MRS Manager as an administrator.

   For details about how to log in to MRS Manager, see **Accessing MRS Manager**.
2. Choose **Cluster** > **Services** > **Yarn**. Click **Configurations** and then **All Configurations**, search for **yarn.resourcemanager.webapp.pagination.enable**, and check whether its value is **true**.
   - If yes, go to **Step 1.3**.
   - If no, set the parameter to **true** and click **Save** to save the configuration. On the **Dashboard** tab page of YARN, choose **More** > **Restart Service**, verify the identity, and click **OK**. After the service restarts, go to **Step 1.3**.
3. Choose **Cluster** > *Name of the desired cluster* > **Services**, and check whether the YARN service status is normal.

Step 2 Log in to the active management node as user **omm**.

Step 3 Switch the scheduler.

The following switching modes are available:

0: converts the Capacity scheduler configurations into the Superior scheduler configurations and then switches the Capacity scheduler to the Superior scheduler.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1: converts the Capacity scheduler configurations into the Superior scheduler configurations only.

2: switches the Capacity scheduler to the Superior scheduler only.

- Mode **0** is recommended if the cluster environment is simple and the number of tenants is less than 20.

  Run the following command:

  sh ${BIGDATA_HOME}/om-server/om/sbin/switchScheduler.sh -c *Cluster ID* **-m 0**

  You can choose **Cluster**, click the cluster name, and choose **Cluster Properties** on FusionInsight Manager to view the cluster ID.

  ```
  Start to convert Capacity scheduler to Superior Scheduler, clusterId=1
  Start to convert Capacity scheduler configurations to Superior. Please wait...
  Convert configurations successfully.
  Start to switch the Yarn scheduler to Superior. Please wait...
  Switch the Yarn scheduler to Superior successfully.
  ```

- If the cluster environment or tenant information is complex and you need to retain the queue configurations of the Capacity scheduler on the Superior scheduler, it is recommended that you use mode **1** first to convert the Capacity scheduler configurations, check the converted configurations, and then use mode **2** to switch the Capacity scheduler to the Superior scheduler.

  a. Run the following command to convert the Capacity scheduler configurations into the Superior scheduler configurations:

     sh ${BIGDATA_HOME}/om-server/om/sbin/switchScheduler.sh -c *Cluster ID* **-m 1**

     ```
     Start to convert Capacity scheduler to Superior Scheduler, clusterId=1
     Start to convert Capacity scheduler configurations to Superior. Please wait...
     Convert configurations successfully.
     ```

  b. Run the following command to switch the Capacity scheduler to the Superior scheduler:

     sh ${BIGDATA_HOME}/om-server/om/sbin/switchScheduler.sh -c *Cluster ID* **-m 2**

     ```
     Start to convert Capacity scheduler to Superior Scheduler, clusterId=1
     Start to switch the Yarn scheduler to Superior. Please wait...
     Switch the Yarn scheduler to Superior successfully.
     ```

- If you do not need to retain the queue configurations of the Capacity scheduler, use mode **2**.

  a. Log in to the MRS Manager by referring to **Accessing MRS Manager** and delete all tenants except the **default** tenant.

  b. Log in to the MRS Manager by referring to **Accessing MRS Manager** and delete all resource pools except the **default** resource pool.

     Run the following command to switch the Capacity scheduler to the Superior scheduler:

     sh ${BIGDATA_HOME}/om-server/om/sbin/switchScheduler.sh -c *Cluster ID* **-m 2**

     ```
     Start to convert Capacity scheduler to Superior Scheduler, clusterId=1
     Start to switch the Yarn scheduler to Superior. Please wait...
     Switch the Yarn scheduler to Superior successfully.
     ```

You can query the scheduler switching logs on the active management node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- ${BIGDATA_LOG_HOME}/controller/aos/switch_scheduler.log
- ${BIGDATA_LOG_HOME}/controller/aos/aos.log

----End

8.10 Managing MRS Cluster Users

8.10.1 Cluster User Permissions

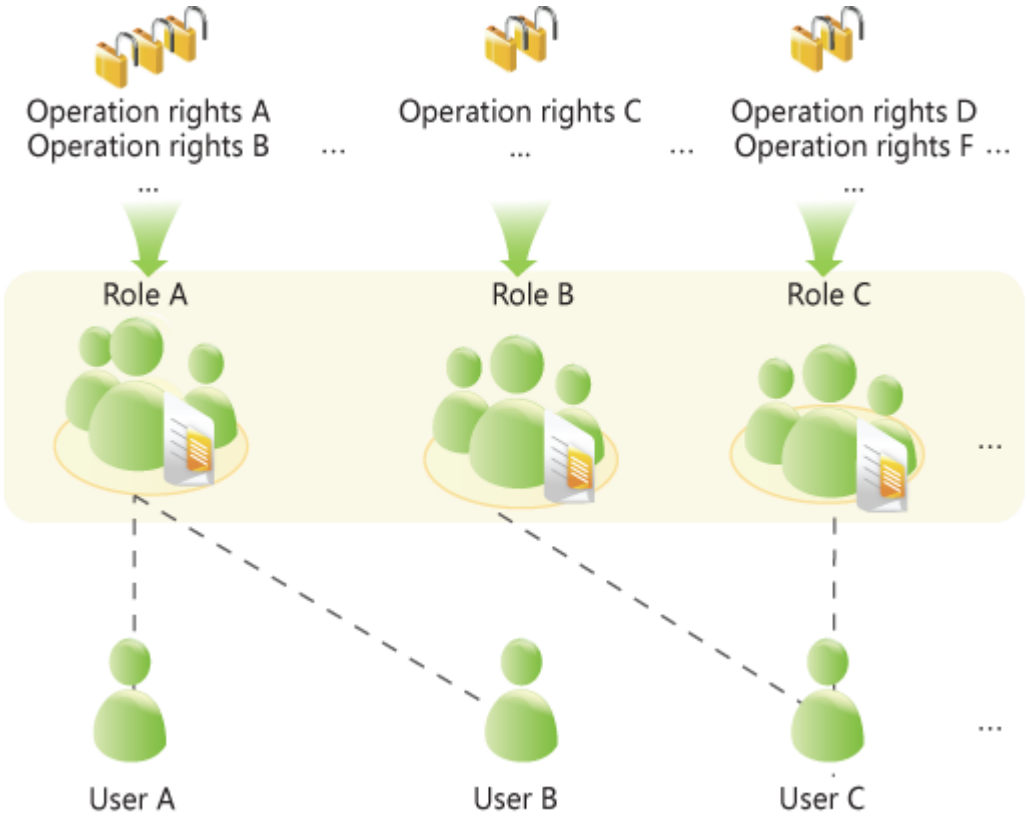
8.10.1.1 MRS Cluster User Permission Model

MRS uses a role-based access control (RBAC) model to manage permissions across the big data system. It consolidates permission management functions from various components into a centralized framework, making permission management easier for administrators. End users do not need to understand the underlying permission mechanisms, simplifying operations and enhancing overall usability and experience.

Role-based Access Control

The right model of MRS consists four parts: users, user groups, roles, and rights.

Figure 8-10 Right model



- **Right**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Right, which is defined by components, allows users to access a certain resource of one component. Different components have different permissions for their resources.

For example:

- HDFS provides read, write, and execute permissions on files.
- HBase provides create, read, and write permissions on tables.

- **Role**

  A role is a collection of permissions associated with various components. Each role may include permissions across multiple components, and multiple roles can share permissions on the same resource within a component.

- **User group**

  A user group is a collection of users. When a user group is associated with a role, users in this group obtain the permissions defined by the role.

  Different user groups can be associated with the same role. A user group can also be associated with no role, and this user group does not have the permissions of any component resources.

  In some components, the system grants related permissions to specific user groups by default.

- **User**

  A user is a visitor to the system. Each user has the permissions on the user group and role associated with the user. Users need to be added to the user group or associated with roles to obtain the corresponding permissions.

Policy-based Access Control

The Ranger component uses policy-based access control (PBAC) to manage permissions and implement fine-grained data access control on components such as HDFS, Hive, and HBase.

A component supports only one access control mechanism. After a Ranger permission control policy is enabled for a component, the permission on the component in the role created on FusionInsight Manager becomes invalid (ACL rules of HDFS and YARN still take effect). You need to add a policy on the Ranger management page to grant permissions to resources.

The Ranger right model consists of multiple right policies. A right policy consists of the following parts:

- Resource

  Resources are provided by components and can be accessed by users, such as HDFS files or folders, queues in Yarn, and databases, tables, and columns in Hive.

- User

  Users are visitors to the system. Each user obtains corresponding permissions after being associated with related policies. Information about users, user groups, and roles in the LDAP is periodically synchronized to the Ranger.

- Permission

  In a policy, you can configure various access conditions for resources, such as file read and write, permission conditions, rejection conditions, and exception conditions.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Example Access Control Scenarios

MRS adopts the Lightweight Directory Access Protocol (LDAP) to store data of users and user groups. Information about role definitions is stored in the relational database and the mapping between roles and permissions is saved in components.

MRS uses Kerberos for unified authentication.

The user permission verification process is as follows:

1. A client (a user terminal or MRS component service) invokes the MRS authentication interface.
2. MRS uses the login username and password for Kerberos authentication.
3. If the authentication succeeds, the client sends a request for accessing the server (an MRS component service).
4. The server finds the user group and role to which the login user belongs.
5. The server obtains all permissions of the user group and the role.
6. The server checks whether the client has the permission to access the resources it applies for.

Example (RBAC):

There are three files in HDFS, that is, fileA, fileB, and fileC.

- roleA has read and write permissions on fileA, and roleB has the read permission on fileB.
- groupA is bound to roleA, and groupB is bound to roleB.
- userA belongs to groupA and roleB, and userB belongs to groupB.

When userA successfully logs in to the system and accesses the HDFS:

1. HDFS obtains the role (roleB) to which userA is bound.
2. HDFS also obtains the role (roleA) to which the user group of userA is bound.
3. In this case, userA has all the permissions of roleA and roleB.
4. As a result, userA has read and write permissions on fileA, has the read permission on fileB, and has no permission on fileC.

Similarly, when userB successfully logs in to the system and accesses the HDFS:

1. userB only has the permissions of roleB.
2. As a result, userB has the read permission on fileB, and has no permission on fileA and fileC.

8.10.1.2 MRS Cluster User Identity Authentication Policy

The big data platform performs user identity authentication to prevent invalid users from accessing the cluster. The cluster provides authentication capabilities in both security mode and normal mode.

Security Mode

The clusters in security mode use the Kerberos authentication protocol for security authentication. The Kerberos protocol supports mutual authentication between

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

clients and servers. This eliminates the risks incurred by sending user credentials over the network for simulated authentication. In clusters, KrbServer provides the Kerberos authentication support.

Kerberos user object

In the Kerberos protocol, each user object is a principal. A complete principal consists of username and domain name. In O&M or application development scenarios, the user identity must be verified before a client connects to a server. Users for O&M and service operations are classified into human-machine and machine-machine users. The password of human-machine users is manually configured, while the password of machine-machine users is generated by the system randomly.

Kerberos authentication

Kerberos supports password and keytab authentication. The validity period of authentication is 24 hours by default.

- Password authentication: User identity is verified by entering the correct password. This mode mainly used in O&M scenarios where human-machine users are used. The configuration command is **kinit** *Username*.
- Keytab authentication: Keytab files contain users' principal and encrypted credential information. When keytab files are used for authentication, the system automatically uses encrypted credential information to perform authentication and the user password does not need to be entered. This mode is mainly used in component application development scenarios where machine-machine users are used. Keytab authentication can also be configured using the **kinit** command.

Normal Mode

Different components in a normal cluster use the native open-source authentication mode and do not support the **kinit** authentication command. FusionInsight Manager (including DBService, KrbServer, and LdapServer) uses the username and password for authentication. **Table 8-60** lists the authentication modes used by components.

Table 8-60 Component authentication modes

Service	Authentication Mode
IoTDB	Simple authentication
CDL	No authentication
ClickHouse	Simple authentication
Flume	No authentication
HBase	• Web UI: No authentication • Client: simple authentication
HDFS	• Web UI: no authentication • Client: simple authentication

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Authentication Mode
Hive	Simple authentication
Hue	Username and password authentication
Kafka	No authentication
Loader	• Web UI: username and password authentication • Client: no authentication
MapReduce	• Web UI: no authentication • Client: no authentication
Oozie	• Web UI: username and password authentication • Client: simple authentication
Spark2x	• Web UI: no authentication • Client: simple authentication
Storm	No authentication
YARN	• Web UI: no authentication • Client: simple authentication
ZooKeeper	Simple authentication

The authentication modes are as follows:

- Simple authentication: When the client connects to the server, the client automatically authenticates the user (for example, the OS user **root** or **omm**) by default. The authentication is imperceptible to the administrator or service user, which does not require **kinit**.
- Username and password authentication: Use the username and password of human-machine users in the cluster for authentication.
- No authentication: Any user can access the server by default.

8.10.1.3 MRS Cluster User Permission Authentication Policy

Security Mode

After a user is authenticated by the big data platform, the system determines whether to verify the user's permission based on the actual permission management configuration to ensure that the user has limited or all permission on resources. If the user does not have the permission for accessing cluster resources, the system administrator must grant the required permission to the user. Otherwise, the user fails to access the resources. The cluster provides permission verification capabilities in both security mode and normal mode. The specific permission items of the components are the same in the two modes.

By default, the Ranger service is installed and Ranger authentication is enabled for a newly installed cluster in security mode. You can set fine-grained security access

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

policies for accessing component resources through the permission plug-in of the component. If Ranger authentication is not required, administrators can manually disable it on the service page. After Ranger authentication is disabled, the system continues to perform permission control based on the role model of FusionInsight Manager when accessing component resources.

In a cluster in security mode, the following components support Ranger authentication: HDFS, YARN, Kafka, Hive, HBase, Storm, Impala, CDL, and Spark2x.

For a cluster upgraded from an earlier version, Ranger authentication is not used by default when users access component resources. The administrator can manually enable Ranger authentication after installing Ranger.

By default, all components in the cluster of the security edition authenticate access. The authentication function cannot be disabled.

Normal Mode

Different components in a normal cluster use their own native open-source authentication behavior. **Table 8-61** lists detailed permission verification modes.

In a normal cluster, Ranger supports permission control on component resources based on OS users. The following components support Ranger authentication: HBase, HDFS, Hive, Spark2x, and YARN.

Table 8-61 Component permission verification modes in normal clusters

Service	Permission Verification	Permission Verification Enabling and Disabling
IoTDB	Required	Not supported
ClickHouse	Required	Not supported
Flume	Not required	Not supported
HBase	Not required	Supported
HDFS	Required	Supported
Hive	Not required	Not supported
Hue	Not required	Not supported
Kafka	Not required	Not supported
Loader	Not required	Not supported
MapReduce	Not required	Not supported
Oozie	Required	Not supported
Spark2x	Not required	Not supported
Storm	Not required	Not supported
YARN	Not required	Supported
ZooKeeper	Required	Supported

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

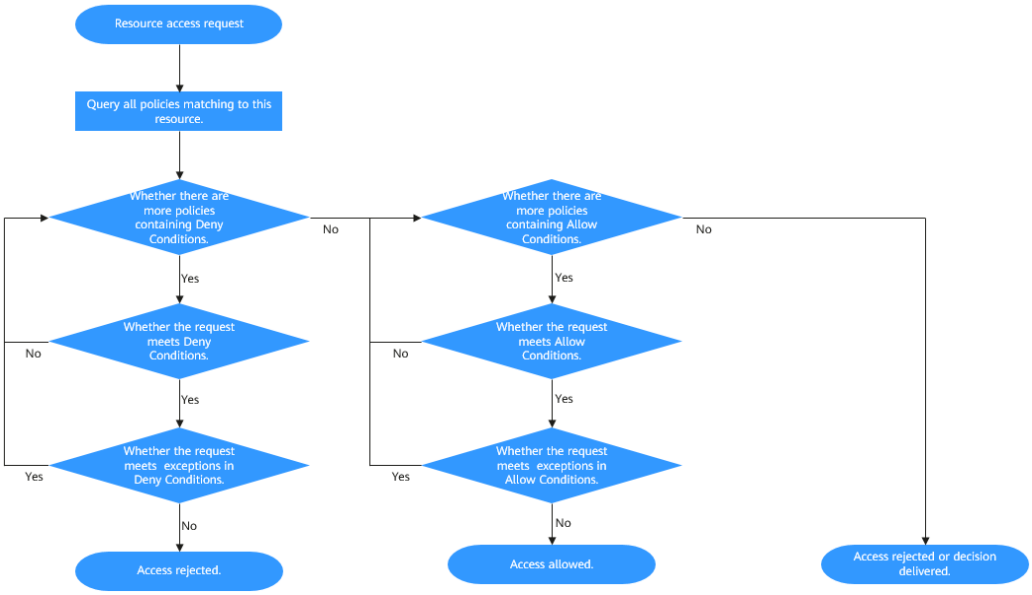
Service	Permission Verification	Permission Verification Enabling and Disabling
CDL	Not required	Not supported

Condition Priorities of the Ranger Permission Policy

When configuring a permission policy for a resource, you can configure Allow Conditions, Exclude from Allow Conditions, Deny Conditions, and Exclude from Deny Conditions for the resource, to meet unexpected requirements in different scenarios.

The priorities of different conditions are listed in descending order: Exclude from Deny Conditions > Deny Conditions > Exclude from Allow Conditions > Allow Conditions

The following figure shows the process of determining condition priorities. If the component resource request does not match the permission policy in Ranger, the system rejects the access by default. However, for HDFS and Yarn, the system delivers the decision to the access control layer of the component for determination.



For example, if you want to grant the read and write permissions of the **FileA** folder to the **groupA** user group, but the user in the group is not **UserA**, you can add an allowed condition and an exception condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.1.4 Default Permissions of the MRS Cluster

Roles

Default Role	Description
Manager_administrator	Manager administrator who has all permissions for Manager. Manager administrators can create first-level tenants, create and modify user groups, and specify user permissions.
Manager_operator	Manager operator who has all the permissions on the **Homepage**, **Cluster**, **Hosts**, and **O&M** tabs.
Manager_auditor	Manager auditor who has all permissions on the **Audit** tab page. Manager auditors can view and manage Manager system audit logs.
Manager_viewer	Manager viewer, who has permissions to view the **Homepage**, **Cluster**, **Hosts**, **Alarms** and **Events** in the O&M alarm module, **Tenant Resources** (available in MRS 3.5.0 and later), and **System** pages, and download clients. (Only MRS 3.2.0 or later supports client download.)
Manager_tenant	Manager tenant administrator. This role can create and manage sub-tenants for the non-leaf tenants to which the current user belongs. It has the permission to view alarms and events on **O&M** > **Alarm** page. **NOTE** MRS 3.5.0 and later clusters: Assigning this role to a user with both **Manager_viewer** and **Manager_tenant** roles overrides the tenant management permission of the **Manager_tenant** role. The user can then only view tenant resources.
System_administrator	System administrator, this role has Manager system administrator rights and all services administrator rights.
default	This role is the default role created for the **default** tenant. It has the management permissions on the YARN component and the default queue. The default role of the default tenant that is not the first cluster to be installed is **c***<cluster ID>*_**default**.
Manager_administrator_180	FusionInsight Manager system administrator group. Internal system user group, which is used only between components.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Default Role	Description
Manager_auditor_181	FusionInsight Manager system auditor group. Internal system user group, which is used only between components.
Manager_operator_182	FusionInsight Manager system operator group. Internal system user group, which is used only between components.
Manager_viewer_183	FusionInsight Manager system viewer group. Internal system user group, which is used only between components.
System_administrator_186	System administrator group. Internal system user group, which is used only between components.
Manager_tenant_187	Tenant system user group. Internal system user group, which is used only between components.
default_1000	User group created for tenants. Internal system user group, which is used only between components.

User Groups

Type	Default User Group	Description
Default cluster user groups	hadoop	Users added to this user group have the permission to submit tasks to all YARN queues.
	hadoopmanager	Users added to this user group can have the O&M manager rights of HDFS and YARN. The O&M manager of HDFS can access the NameNode web UI and perform active-standby switchover. The O&M manager of YARN can access the ResourceManager web UI, operate NodeManager nodes, refresh queues, and set node labels, but cannot submit tasks.
	hetuadmin	HetuEngine administrator group. Users in this group have the permission to perform operations on HSConsole.
	hive	Common user group. Hive users must belong to this user group.
	iotdbgroup	Users added to this user group have the administrator rights of the IoTDB component.
	kafka	Common Kafka user group. A user in this group can access a topic only when a user in the **kafkaadmin** group grants the read and write permission of the topic to the user.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Default User Group	Description
	kafkaadmin	Kafka administrator group. Users added to this group have the permissions to create, delete, authorize, as well as read from and write data to all topics.
	kafkasuperuser	Topic read/write user group of Kafka. Users added to this group have permissions to read data from and write data to all topics.
	cdladmin	CDL administrator group. Only users in this group can access CDL APIs.
	cdl	Common user group of CDL. Users in this group can create and query CDL jobs.
	storm	Storm common user group. Users added to this group have the permissions to submit topologies and manage their own topologies.
	stormadmin	Storm administrator user group. Users added to this group have the permissions to submit topologies and manage their own topologies.
	supergroup	Users added to this user group have the administrator rights of HBase, HDFS, and YARN and can use Hive.
	yarnviewgroup	Read-only user group of YARN tasks. Users in this user group can have the view permission on YARN and MapReduce tasks.
	check_sec_ldap	User group for testing whether the active LDAP works properly. This user group is generated randomly in a test and automatically deleted after the test is complete. This is an internal system user group used only between components.
	compcommon	System internal group for accessing cluster system resources. All system users and system running users are added to this user group by default.
Default OS user groups	wheel	Primary group of the internal system running user **omm**.
	ficommon	System common group that corresponds to **compcommon** for accessing cluster common resource files stored in the OS.

Users

For details, see **MRS Cluster User Accounts**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.1.5 Synchronizing IAM Users to MRS

IAM user synchronization is to synchronize IAM users bound with MRS policies to the MRS system and create accounts with the same usernames but different passwords as the IAM users. Then, you can use an IAM username (the password needs to be reset by user **admin** of Manager) to log in to Manager for cluster management, and submit jobs on the GUI in a cluster with Kerberos authentication enabled.

Table 8-62 compares IAM users' permission policies and the synchronized users' permissions on MRS. For details about the default permissions on Manager, see **Default Permissions of the MRS Cluster**.

Table 8-62 Policy and permission mapping after synchronization

Policy Type	IAM Policy	User's Default Permissions on MRS After Synchronization	Have Permission to Perform the Synchronization	Have Permission to Submit Jobs
Fine-grained	MRS ReadOnlyAccess	Manager_viewer	No	No
	MRS CommonOperations	• Manager_viewer • default • launcher-job	No	Yes
	MRS FullAccess	• Manager_administrator • Manager_auditor • Manager_operator • Manager_tenant • Manager_viewer • System_administrator • default • launcher-job	Yes	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Policy Type	IAM Policy	User's Default Permissions on MRS After Synchronization	Have Permission to Perform the Synchronization	Have Permission to Submit Jobs
RBAC	MRS Administrator	• Manager_administrator • Manager_auditor • Manager_operator • Manager_tenant • Manager_viewer • System_administrator • default • launcher-job	No	Yes
	Server Administrator, Tenant Guest, and MRS Administrator	• Manager_administrator • Manager_auditor • Manager_operator • Manager_tenant • Manager_viewer • System_administrator • default • launcher-job	Yes	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Policy Type	IAM Policy	User's Default Permissions on MRS After Synchronization	Have Permission to Perform the Synchronization	Have Permission to Submit Jobs
	Tenant Administrator	• Manager_administrator • Manager_auditor • Manager_operator • Manager_tenant • Manager_viewer • System_administrator • default • launcher-job	Yes	Yes
Custom	Custom policy	• Manager_viewer • default • launcher-job	• If custom policies use RBAC policies as a template, refer to the RBAC policies. • If custom policies use fine-grained policies as a template, refer to the fine-grained policies. The fine-grained policies are recommended.	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For easier permission management, use fine-grained policies instead of RBAC. Within fine-grained policies, the Deny action takes precedence over all other actions.

- To synchronize IAM users, a user must either have the Tenant Administrator role, or simultaneously hold the Server Administrator, Tenant Guest, and MRS Administrator roles.
- A user with the **action:mrs:cluster:syncUser** policy can synchronize IAM users.

Constraints on IAM User Synchronization

Because of restrictions in open-source components, some IAM users cannot be synchronized. For example, usernames with special characters such as periods (.) or at signs (@), or usernames with only one character, are not supported. To ensure successful synchronization of IAM users, the usernames must conform to the naming conventions defined by MRS Manager.

- In MRS 3.*x* and later versions, usernames must contain 2 to 64 characters. Valid characters include digits, case-insensitive letters, underscores (_), hyphens (-), periods (.), and spaces. Usernames must not begin with a hyphen (-) or contain consecutive periods (..). Each username must be unique within the system or operating system.
- In versions earlier than MRS 3.*x*, usernames must contain 3 to 32 characters. Valid characters include letters, digits, spaces, hyphens (-), and underscores (_). Usernames must not begin with a digit, nor duplicate existing operating system accounts such as **root** or **omm**.

Synchronizing IAM Users

Step 1 Create a user and authorize the user to use MRS. For details, see **Creating an IAM User and Granting MRS Permissions**.

Step 2 Log in to the MRS management console and create a cluster. For details, see **Creating a Custom MRS Cluster**.

Step 3 In the navigation pane on the left, choose **Active Clusters**. Click a cluster name to go to the cluster details page.

Step 4 On the **Dashboard** tab page, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.

Step 5 In the **IAM User Sync** dialog box, search for the user group to which the IAM user to be synchronized belongs and click the user group name. In the **User** column, select the desired IAM user and click **Synchronize**.

- To synchronize all users, select **All** in the **To Synchronize** tab.
- If you select only user groups, users will not be synchronized. To synchronize users, specific usernames within those groups must be selected.
- All user groups are displayed in the **IAM User Sync** window. User groups and usernames that cannot be selected cannot be synchronized.

Step 6 After a synchronization request is sent, choose **Operation Logs** in the navigation tree on the left of the MRS console to check whether the synchronization is successful. For details about the logs, see **Viewing MRS Operation Logs**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 After the synchronization is successful, use the user synchronized with IAM to perform subsequent operations.

NOTE

- When the policy of the user group to which the IAM user belongs changes from **MRS ReadOnlyAccess** to **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator**, wait for 5 minutes until the new policy takes effect after the synchronization is complete because the **SSSD** (System Security Services Daemon) cache of cluster nodes needs time to be updated. Then, submit a job. Otherwise, the job may fail to be submitted.
- When the policy of the user group to which the IAM user belongs changes from **MRS CommonOperations**, **MRS FullAccess**, or **MRS Administrator** to **MRS ReadOnlyAccess**, wait for 5 minutes until the new policy takes effect after the synchronization is complete because the **SSSD** cache of cluster nodes needs time to be updated.
- After you click **Synchronize** on the right side of **IAM User Sync**, the cluster details page is blank for a short time, because user data is being synchronized. The page will be properly displayed after the data synchronization is complete.

- Submitting jobs in a security cluster: Users can submit jobs using the job management function on the GUI in the security cluster. For details, see **Running a MapReduce Job**.
- All tabs are displayed on the cluster details page, including **Components**, **Tenants**, and **Backups & Restorations**.
- Logging in to Manager

  a. Log in to Manager as user **admin**. For details, see **Accessing MRS Manager**.

  b. Initialize the password of the user synchronized with IAM. For details, see **Initializing MRS Cluster User Passwords**.

  c. Modify the role bound to the user group to which the user belongs to control user permissions on Manager. For details, see **Managing MRS Cluster User Groups**. For details about how to create and modify a role, see **Creating a Role**. After the component role bound to the user group to which the user belongs is modified, it takes some time for the role permissions to take effect.

  d. Log in to Manager using the user synchronized with IAM and the password after the initialization in **Step 7.b**.

  If an IAM user's permission changes, go to **Step 4** to perform second synchronization. After the second synchronization, a system user's effective permissions are determined as the union of the permissions defined in the IAM system policy and those granted through roles added by the user in MRS Manager. After the second synchronization, the effective permissions of a custom user are determined solely by the permissions configured on MRS Manager.

  – System user: An IAM user is considered a system user if every user group to which they belong is bound to system policies (including RBAC and fine-grained policies).

  – Custom user: An IAM user is considered a custom user if the user group to which they belong is bound to any custom policy.

Step 8 Undo IAM user synchronization.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

To undo the synchronization of an IAM user, select the user in the **User** column in the **Synchronized** tab and click **Undo Sync**.

To undo the synchronization of all users in an IAM user group, select the user group in the **User Group** column in the **Synchronized** tab and click **Undo Sync**.

----End

8.10.2 MRS Cluster User Accounts

This section describes information about default users in MRS clusters.

Account List (MRS 3.x and Later Versions)

- **User types**

  The MRS cluster provides the following three types of users. The system administrator needs to periodically change the passwords. It is not recommended to use the default passwords.

User Type	Description
System users	• User created on FusionInsight Manager for O&M and service scenarios. There are two types of users: – **Human-machine** user: used in scenarios such as FusionInsight Manager O&M and operations on a component client. When creating a user of this type, you need to set password and confirm password by referring to **Creating an MRS Cluster User**. – **Machine-machine** user: used for system application development. • User who runs OMS processes
Internal system users	Internal user to perform Kerberos authentication, process communications, save user group information, and associate user permissions. It is recommended that internal system users not be used in O&M scenarios. Operations can be performed as user **admin** or another user created by the system administrator based on service requirements.
Database users	• User who manages OMS database and accesses data • User who runs service components (Hue, Hive, HetuEngine, Loader, Oozie, Ranger, JobGateway, and DBService) in the database.

- **System user**
  - User **root** of the OS is required, the password of user **root** on all nodes must be the same.
  - User **ldap** of the OS is required. Do not delete this account. Otherwise, the cluster may not work properly. The OS administrator is responsible for maintaining password management policies.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Username	Initial Password	Description	Password Change Method
System administrator	admin	User-defined password	FusionInsight Manager administrator. By default, user **admin** does not have management permission on other components. For example, when accessing the native UI of a component, the user cannot view complete component information due to insufficient management permissions on the component.	For details, see **Changing or Resetting the Password for User admin of an MRS Cluster**.
Node OS user	ommdba	Random password	User that creates the system database. This user is an OS user generated on the management node and does not require a unified password. This account cannot be used for remote login.	For details, see **Changing the Passwords for OS Users of an MRS Cluster Node**.
	omm	Random password	Internal running user of the system. This user is an OS user generated on all nodes and does not require a unified password.	

- **Internal system user**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
Kerberos administrator	kadmin/admin	Admin@123	Used to add, delete, modify, and query user accounts on Kerberos.	For details, see **Changing the Password for the Kerberos Administrator of an MRS Cluster**.
OMS Kerberos administrator	kadmin/admin	Admin@123	Used to add, delete, modify, and query user accounts on OMS Kerberos.	For details, see **Changing the Password of the OMS Kerberos Administrator**.
LDAP administrator	cn=root,dc=hadoop,dc=com	• Versions earlier than MRS 3.1.2: LdapChangeMe@123 • MRS 3.1.2 or later: randomly generated by the system	Used to add, delete, modify, and query the user account information on LDAP.	• For versions earlier than MRS 3.1.2, see **Changing the Password for a Regular LDAP User of an MRS Cluster**. • For MRS 3.1.2 or later, see **Modifying the OMS Service Configuration**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
OMS LDAP administrator	cn=root,dc=hadoop,dc=com	• Versions earlier than MRS 3.1.2: LdapChangeMe@123 • MRS 3.1.2 or later: randomly generated by the system	Used to add, delete, modify, and query the user account information on OMS LDAP.	
LDAP user	cn=pg_search_dn,ou=Users,dc=hadoop,dc=com	Randomly generated by the system	Used to query information about users and user groups on LDAP.	
OMS LDAP user	cn=pg_search_dn,ou=Users,dc=hadoop,dc=com	Randomly generated by the system	Used to query information about users and user groups on OMS LDAP.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
LDAP administrator account	cn=krbkdc,ou=Users,dc=hadoop,dc=com	• Versions earlier than MRS 3.1.2: LdapChangeMe@123 • MRS 3.1.2 or later: randomly generated by the system	Used to query Kerberos component authentication account information.	• For versions earlier than MRS 3.1.2, see **Changing the LDAP Administrator Password for an MRS Cluster**. • For MRS 3.1.2 or later, see **Modifying the OMS Service Configuration**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	cn=krbadmin,ou=Users,dc=hadoop,dc=com	• Versions earlier than MRS 3.1.2: LdapChangeMe@123 • MRS 3.1.2 or later: randomly generated by the system	Used to add, delete, modify, and query Kerberos component authentication account information.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
Component running user	hdfs	Hdfs@123	This user is the HDFS system administrator and has the following permissions: 1. File system operation permissions: • Views, modifies, and creates files. • Views and creates directories. • Views and modifies the groups where files belong. • Views and sets disk quotas for users. 2. HDFS management operation permissions: • Views the web UI status. • Views and sets the active and standby HDFS status. • Enters and exits the HDFS in security mode. • Checks the HDFS file system. 3. Logs in to the FTP service page.	For details, see **Changing the Passwords for MRS Cluster Component Running Users**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	hbase	Hbase@123	This user is the HBase and HBase1 to HBase4 system administrator and has the following permissions: • Cluster management permission: Performs **Enable** and **Disable** operations on tables to trigger MajorCompact and ACL operations. • Grants and revokes permissions, and shuts down the cluster. • Table management permission: Creates, modifies, and deletes tables. • Data management permission: Reads data in tables, column families, and columns. • Logs in to the HMaster web UI. • Logs in to the FTP service page.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	cdl	CDCUser123!	System administrator of the CDL Currently, CDL does not involve user permissions.	

User Type	Default User	Initial Password	Description	Password Change Method
	iotdb	Iotdb@123	This user is the IoTDB system administrator and has the following user permissions: 1. IoTDB administrator permissions: ● Creates or deletes a storage group. ● Uses TTL. 2. IoTDB data operation permissions: ● Creates, modifies, and deletes a time sequence. ● Writes, reads, and deletes data in a time sequence. 3. Views user or role permission information. 4. Grants or revokes permissions for a user or role. In a normal cluster, the IoTDB service retains the features of open-source versions. The default username is **root**, and the default password is **root** (for MRS 3.3.0 or later, the default password is **Iotdb@123**). This user is an	

User Type	Default User	Initial Password	Description	Password Change Method
			administrator and has all permissions, which cannot be assigned, revoked, or deleted.	
	mapred	Mapred@123	This user is the MapReduce system administrator and has the following permissions: • Submits, stops, and views the MapReduce tasks. • Modifies the Yarn configuration parameters. • Logs in to the FTP service page. • Logs in to the Yarn web UI.	
	zookeeper	ZooKeeper@123	This user is the ZooKeeper system administrator and has the following permissions: • Adds, deletes, modifies, and queries all nodes in ZooKeeper. • Modifies and queries quotas of all nodes in ZooKeeper.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	rangeradmin	Rangeradmin@123	This user has the Ranger system management permissions and user permissions: • Ranger web UI management permission • Management permission of each component that uses Ranger authentication	
	rangerauditor	Rangerauditor@123	Default audit user of the Ranger system.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	hive	Hive@123	This user is the Hive system administrator and has the following permissions: 1. Hive administrator permissions: ● Creates, deletes, and modifies a database. ● Creates, queries, modifies, and deletes a table. ● Queries, inserts, and uploads data. 2. HDFS file operation permissions: ● Views, modifies, and creates files. ● Views and creates directories. ● Views and modifies the groups where files belong. 3. Submits and stops the MapReduce tasks. 4. Ranger policy management permission	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	kafka	Kafka@123	This user is the Kafka system administrator and has the following permissions: • Creates, deletes, produces, and consumes the topic; modifies the topic configuration. • Controls the cluster metadata, modifies the configuration, migrates the replica, elects the leader, and manages ACL. • Submits, queries, and deletes the consumer group offset. • Queries the delegation token. • Queries and submits the transaction.	
	storm	Admin@123	Storm system administrator User permission: Submits Storm tasks.	
	rangerusersync	Randomly generated by the system	Synchronizes users and internal users of user groups.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	rangertagsync	Randomly generated by the system	Internal user for synchronizing tags.	
	rangerobs/hadoop.*<System domain name>*	Randomly generated by the system	System administrator used by Guardian to access Ranger	
	jobserver	Randomly generated by the system	JobGateway system administrator, who has the following permissions: 1. HDFS file operation permissions: • Views, modifies, and creates files. • Views and creates directories. • Views and modifies the groups where files belong. 2. Manager administrator permission	
	HTTP/_HOST	Randomly generated by the system	Internal user of the JobGateway service, which is used for Kerberos authentication of the HTTP service	
	oms/manager	Randomly generated by the system	Controller and NodeAgent authentication user. The user has the permission on the **supergroup** group.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	backup/ manager	Randomly generated by the system	User for running backup and restoration tasks. The user has the permission on the **supergroup**, **wheel**, and **ficommon** groups. After cross-system mutual trust is configured, the user has the permission to access data in the HDFS, HBase, Hive, and ZooKeeper systems.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	hdfs/ hadoop.<*System domain name*>	Randomly generated by the system	This user is used to start the HDFS and has the following permissions: 1. File system operation permissions: • Views, modifies, and creates files. • Views and creates directories. • Views and modifies the groups where files belong. • Views and sets disk quotas for users. 2. HDFS management operation permissions: • Views the web UI status. • Views and sets the active and standby HDFS status. • Enters and exits the HDFS in security mode. • Checks the HDFS file system. 3. Logs in to the FTP service page.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	hetuserver/hadoop.*<System domain name>*	Randomly generated by the system	This user is used to start HetuEngine and has the following permissions: • Accesses KrbServer and HDFS files in the cluster from HetuEngine. • Used for communication between HetuEngine internal nodes.	
	mapred/hadoop.*<System domain name>*	Randomly generated by the system	This user is used to start the MapReduce and has the following permissions: • Submits, stops, and views the MapReduce tasks. • Modifies the Yarn configuration parameters. • Logs in to the FTP service page. • Logs in to the Yarn web UI.	
	mr_zk/hadoop.*<System domain name>*	Randomly generated by the system	Used for MapReduce to access ZooKeeper.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	hbase/hadoop.*<System domain name>*	Randomly generated by the system	User for the authentication between internal components during the HBase system startup.	
	hbase/zkclient.*<System domain name>*	Randomly generated by the system	User for HBase to perform ZooKeeper authentication in a security mode cluster.	
	thrift/hadoop.*<System domain name>*	Randomly generated by the system	ThriftServer system startup user.	
	thrift/*<hostname>*	Randomly generated by the system	User for the ThriftServer system to access HBase. This user has the read, write, execution, creation, and administration permission on all NameSpaces and tables of HBase. *<hostname>* indicates the name of the host where the ThriftServer node is installed in the cluster.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	hive/hadoop.*<System domain name>*	Randomly generated by the system	User for the authentication between internal components during the Hive system startup. The user permissions are as follows: 1. Hive administrator permissions: ● Creates, deletes, and modifies a database. ● Creates, queries, modifies, and deletes a table. ● Queries, inserts, and uploads data. 2. HDFS file operation permissions: ● Views, modifies, and creates files. ● Views and creates directories. ● Views and modifies the groups where files belong. 3. Submits and stops the MapReduce tasks.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	loader/hadoop.*<System domain name>*	Randomly generated by the system	User for Loader system startup and Kerberos authentication	
	HTTP/*<hostname>*	Randomly generated by the system	Used to connect to the HTTP interface of each component. *<hostname>* indicates the host name of a node in the cluster.	
	hue	Randomly generated by the system	User for Hue system startup, Kerberos authentication, and HDFS and Hive access	
	flume	Randomly generated by the system	User for Flume system startup and HDFS and Kafka access. The user has read and write permission of the HDFS directory **/flume**.	
	flume_server	Randomly generated by the system	User for Flume system startup and HDFS and Kafka access. The user has read and write permission of the HDFS directory **/flume**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	spark2x/hadoop.*<System domain name>*	Randomly generated by the system	This user is the Spark2x system administrator and has the following user permissions: 1. Starts the Spark2x service. 2. Submits Spark2x tasks.	
	spark_zk/hadoop.*<System domain name>*	Randomly generated by the system	Used for Spark2x to access ZooKeeper.	
	zookeeper/hadoop.*<System domain name>*	Randomly generated by the system	ZooKeeper system startup user.	
	zkcli/hadoop.*<System domain name>*	Randomly generated by the system	ZooKeeper server login user.	
	oozie	Randomly generated by the system	User for Oozie system startup and Kerberos authentication.	
	kafka/hadoop.*<System domain name>*	Randomly generated by the system	Used for security authentication of Kafka.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	storm/hadoop.*<System domain name>*	Randomly generated by the system	Storm system startup user.	
	storm_zk/hadoop.*<System domain name>*	Randomly generated by the system	Used for the Worker process to access ZooKeeper.	
	flink/hadoop.*<System domain name>*	Randomly generated by the system	Internal user of the Flink service.	
	check_ker_M	Randomly generated by the system	User who performs a system internal test about whether the Kerberos service is normal.	
	cdl/hadoop.*<System domain name>*	Randomly generated by the system	Internal user of the CDL service.	
	clickhouse/hadoop.*<System domain name>*	Randomly generated by the system	Used for security authentication of ClickHouse. This user is an internal user and can be used only in the cluster.	
	default	None	ClickHouse internal user, which is an administrator user that can be used only in non-security mode.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	rangeradmin/ hadoop.*<System domain name>*	Randomly generated by the system	Ranger system startup user, which is used for authentication between internal components.	
	tez	Randomly generated by the system	User for TezUI system startup, Kerberos authentication, and access to Yarn	
	K/M	Randomly generated by the system	Kerberos internal functional user. It cannot be deleted, and its password cannot be changed. This internal account can only be used on nodes where Kerberos service is installed.	None
	kadmin/ changepw	Randomly generated by the system		
	kadmin/ history	Randomly generated by the system		
	krbtgt*<System domain name>*	Randomly generated by the system		

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
LDAP user	admin	None	FusionInsight Manager administrator. The primary group is **compcommon**, which does not have the group permission but has the permission of the **Manager_administrator** role.	The LDAP user cannot log in to the system, and the password cannot be changed.
	backup		The primary group is **compcommon**.	
	backup/ manager		The primary group is **compcommon**.	
	oms		The primary group is **compcommon**.	
	oms/ manager		The primary group is **compcommon**.	
	clientregister		The primary group is **compcommon**.	
	zookeeper		The primary group is **hadoop**.	
	zookeeper / hadoop.*<System domain name>*		The primary group is **hadoop**.	
	zkcli		The primary group is **hadoop**.	
	zkcli/ hadoop.*<System domain name>*		The primary group is **hadoop**.	
	flume		The primary group is **hadoop**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	flume_server		The primary group is **hadoop**.	
	hdfs		The primary group is **hadoop**.	
	hdfs/hadoop.*<System domain name>*		The primary group is **hadoop**.	
	mapred		The primary group is **hadoop**.	
	mapred/hadoop.*<System domain name>*		The primary group is **hadoop**.	
	mr_zk		The primary group is **hadoop**.	
	mr_zk/hadoop.*<System domain name>*		The primary group is **hadoop**.	
	hue		The primary group is **supergroup**.	
	hive		The primary group is **hive**.	
	hive/hadoop.*<System domain name>*		The primary group is **hive**.	
	hbase		The primary group is **hadoop**.	
	hbase/hadoop.*<System domain name>*		The primary group is **hadoop**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	thrift		The primary group is **hadoop**.	
	thrift/hadoop.*<System domain name>*		The primary group is **hadoop**.	
	oozie		The primary group is **hadoop**.	
	hbase/zkclient.*<System domain name>*		The primary group is **hadoop**.	
	loader		The primary group is **hadoop**.	
	loader/hadoop.*<System domain name>*		The primary group is **hadoop**.	
	spark2x		The primary group is **hadoop**.	
	spark2x/hadoop.*<System domain name>*		The primary group is **hadoop**.	
	spark_zk		The primary group is **hadoop**.	
	kafka		The primary group is **kafkaadmin**.	
	kafka/hadoop.*<System domain name>*		The primary group is **kafkaadmin**.	
	storm		The primary group is **stormadmin**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description	Password Change Method
	storm/hadoop.*<System domain name>*		The primary group is **stormadmin**.	
	storm_zk		The primary group is **storm**.	
	storm_zk/hadoop.*<System domain name>*		The primary group is **storm**.	
	kms/hadoop		The primary group is **kmsadmin**.	
	knox		The primary group is **compcommon**.	
	executor		The primary group is **compcommon**.	
	rangeradmin		The primary group is **supergroup**.	
	rangeradmin/hadoop.*<System domain name>*		The primary group is **supergroup**.	
	rangerusersync		The primary group is **supergroup**.	
	rangertagsync		The primary group is **supergroup**.	
	rangeraudit or		The primary group is **compcommon**.	
	jobserver		The primary group is **compcommon**.	

Log in to FusionInsight Manager, choose **System** > **Permission** > **Domain and Mutual Trust**, and check the value of **Local Domain**. In the preceding table, all letters in the system domain name contained in the username of the system internal user are lowercase letters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, if **Local Domain** is set to **9427068F-6EFA-4833-B43E-60CB641E5B6C.COM**, the username of default HDFS startup user is **hdfs/hadoop.9427068f-6efa-4833-b43e-60cb641e5b6c.com**.

- **Database user**

  The system database users include OMS database users and DBService database users.

Database Type	Default User	Initial Password	Description	Password Change Method
OMS database	ommdba	• Version earlier than MRS 3.2.0: **dbChangeMe@123456** • MR	OMS database administrator who performs maintenance operations, such as creating, starting, and stopping.	For details, see **Changing the Password for the OMS Database Administrator**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		S3.2.0orlaterrandompassword		

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
	omm	• Versionsearlierthan MRS 3.2.0: **ChangeMe@123456** • MRS 3	User for accessing OMS database data	For details, see **Changing the Password for an OMS Database Access User**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		.2.0orlaterrandompassword		

Database Type	Default User	Initial Password	Description	Password Change Method
DBService database	omm	• Version earlier than MRS 3.2.0: **dbserverAdmin@123** • MR	Administrator of the GaussDB database in the DBService component	For details, see **Resetting the Password for User omm in DBService**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		S3.2.0orlater:randompassword.Ifyouneedtoobtainth		

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		epassword,contactttechnicalsupport.		
	compdb user	Random password	(Available in MRS 3.1.2 or later) Administrator of the GaussDB database in the DBService component. It is used in service O&M scenarios. You need to reset the password upon your first login.	For details, see **Changing the Password for User compdbuser of the DBService Database**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
	hetu	Random password	User for HetuEngine to connect to the DBService database **hetumeta**. This user exists only in MRS 3.1.2 or later.	• For versions earlier than MRS 3.1.2, see **Changing the Passwords for Database Users of MRS Cluster Components**. • For MRS 3.1.2 or later, see **Resetting the MRS Component Database User Password**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
	hive	• Versionsearlierthan MRS 3.1.2: HiveUser@ • MRS 3.1.2 or	User for Hive to connect to the DBService database **hivemeta**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		later:randompassword		

Database Type	Default User	Initial Password	Description	Password Change Method
	hue	• Versionsearlierthan MRS 3.1.2: HueUser@123 • MRS 3.1.2	User for Hue to connect to the DBService database **hue**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		orlater:randompassword		

Database Type	Default User	Initial Password	Description	Password Change Method
	sqoop	• Versionsearlierthan MRS 3.1.2: SqoopUser@ • MRS 3.1.2o	User for Loader to connect to the DBService database **sqoop**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		rlatter:randompassword		

Database Type	Default User	Initial Password	Description	Password Change Method
	oozie	• Versionsearlierthan MRS 3.1.2: OozieUser@ • MRS 3.1.2o	User for Oozie to connect to the DBService database **oozie**.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		rlatter:randompassword		

Database Type	Default User	Initial Password	Description	Password Change Method
	rangeradmin	• Versionsearlierthan MRS 3.1.2: Admin12! • MRS 3.1.2 or	User for Ranger to connect to the DBService database.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Database Type	Default User	Initial Password	Description	Password Change Method
		ater:randompassword		
	kafkaui	Random password	User for Kafka UI to connect to the DBService database. This user exists only in MRS 3.1.2 or later.	
	flink	Random password	User for Flink to connect to the DBService database **flinkmeta**. This user exists only in MRS 3.1.2 or later.	
	cdl	Random password	User for CDL to connect to the DBService database **cdl**. This user exists only in MRS 3.2.0 or later.	
	jobgateway	Random password	User for JobGateway to connect to the DBService database **jobmeta**. This user can be used in MRS 3.3.0 and later versions only.	

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Account List (MRS 2.x and Earlier Versions)

- **User type**

  The MRS cluster provides the following three types of users. Periodically change the passwords and do not use the default passwords.

User Type	Description
System users	• User created on Manager for MRS cluster O&M and service scenarios. There are two types of system users: – **Human-machine** user: used for Manager O&M scenarios and component client operation scenarios. – **Machine-machine** user: used for MRS cluster application development scenarios. • User who runs OMS processes
Internal system users	Internal user to perform Kerberos authentication, process communications, save user group information, and associate user permissions. It is recommended that internal system users not be used in O&M scenarios. Operations can be performed as user **admin** or another user created by the system administrator based on service requirements.
Database users	• User who manages OMS database and accesses data • User who runs the database of service components (Hive, Loader, and DBService)

- **System user**

  User **ldap** of the OS is required in the MRS cluster. Do not delete this account. Otherwise, the cluster may not work properly. Operation users are responsible for maintaining password management policies.

  When you change the passwords of users **ommdba** and **omm** for the first time, you must reset them. Change the passwords periodically after retrieving them.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Username	Initial Password	Description
System administrator of the MRS cluster	admin	Specified by the user during the cluster creation.	MRS Manager The user has the following permissions: • Common HDFS and ZooKeeper user permissions. • Permissions to submit and query MapReduce and YARN tasks, manage YARN queues, and access the YARN web UI. • Permissions to submit, query, activate, deactivate, reassign, delete topologies, and operate all topologies of the Storm service. • Permissions to create, delete, authorize, reassign, consume, write, and query topics of the Kafka service.
MRS cluster node OS user	omm	Randomly generated by the system	Internal running user of the MRS cluster system. This user is an OS user generated on all nodes and does not require a unified password.
MRS cluster node OS user	root	The password is set by the user.	User for logging in to the node in the MRS cluster. This user is an OS user generated on all nodes.

- **Internal system user**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

WARNING

- Do not delete the following internal system users. Deleting them may cause the cluster or components to malfunction.
- These users are available only in clusters with Kerberos authentication enabled.

User Type	Default User	Initial Password	Description
Component running user	hdfs	Hdfs@123	This user is the HDFS system administrator and has the following permissions: 1. File system operation permissions: • Views, modifies, and creates files. • Views and creates directories. • Views and modifies the groups where files belong. • Views and sets disk quotas for users. 2. HDFS management operation permissions: • Views the web UI status. • Views and sets the active and standby HDFS status. • Enters and exits the HDFS in security mode. • Checks the HDFS file system.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description
	hbase	Hbase@123	This user is the HBase system administrator and has the following permissions: • Cluster management permission: **Enable** and **Disable** operations on tables to trigger MajorCompact and ACL operations. • Grants and revokes permissions, and shuts down the cluster. • Table management permission: Creates, modifies, and deletes tables. • Data management permission: Reads data in tables, column families, and columns. • Accesses the HBase web UI.
	mapred	Mapred@123	This user is the MapReduce system administrator and has the following permissions: • Submits, stops, and views the MapReduce tasks. • Modifies the Yarn configuration parameters. • Accesses the Yarn and MapReduce web UI.
	spark	Spark@123	This user is the Spark system administrator and has the following permissions: • Accesses the Spark web UI. • Submits Spark tasks.

- **User group information**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Default User Group	Description
supergroup	Primary group of user **admin**, which has no additional permissions in the cluster with Kerberos authentication disabled.
check_sec_ldap	Used to test whether the active LDAP works properly. This user group is generated randomly in a test and automatically deleted after the test is complete. This is an internal system user group used only between components.
Manager_tenant	Tenant system user group, which is an internal system user group used only between components. It is used only in clusters with Kerberos authentication enabled.
System_administrator	MRS cluster system administrator group, which is an internal system user group used only between components. It is used only in clusters with Kerberos authentication enabled.
Manager_viewer	MRS Manager system viewer group, which is an internal system user group used only between components. It is used only in clusters with Kerberos authentication enabled.
Manager_operator	MRS Manager system operator group, which is an internal system user group used only between components. It is used only in clusters with Kerberos authentication enabled.
Manager_auditor	MRS Manager system auditor group, which is an internal system user group used only between components. It is used only in clusters with Kerberos authentication enabled.
Manager_administrator	MRS Manager system administrator group, which is an internal system user group used only between components. It is used only in clusters with Kerberos authentication enabled.
compcommon	MRS cluster internal group, used to access public resources in the cluster. All system users and system running users are added to this user group by default.
default_1000	User group created for tenants. This is an internal system user group used only between components.
launcher-job	MRS internal group, which is used to submit jobs using V2 APIs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Default User Group	Description
hadoop	Users added to this user group have the permission to submit tasks to all YARN queues. This type of user is available only in clusters with Kerberos authentication enabled.
hbase	Common user group. Users added to this user group will not have any additional permission. This type of user is available only in clusters with Kerberos authentication enabled.
hive	Users added to this user group can use Hive. This type of user is available only in clusters with Kerberos authentication enabled.
spark	Common user group. Users added to this user group will not have any additional permission. This type of user is available only in clusters with Kerberos authentication enabled.
kafka	Kafka common user group. Users added to this group need to be granted with read and write permission by users in the **kafkaadmin** group before accessing the desired topics. This type of user is available only in clusters with Kerberos authentication enabled.
kafkasuperuser	Users added to this group have permissions to read data from and write data to all topics. This type of user is available only in clusters with Kerberos authentication enabled.
kafkaadmin	Kafka administrator group. Users added to this group have the permissions to create, delete, authorize, as well as read from and write data to all topics. This type of user is available only in clusters with Kerberos authentication enabled.
storm	Storm common user group. Users added to this group have the permissions to submit topologies and manage their own topologies. This type of user is available only in clusters with Kerberos authentication enabled.
stormadmin	Storm administrator user group. Users added to this group have the permissions to submit topologies and manage their own topologies. This type of user is available only in clusters with Kerberos authentication enabled.
opentsdb	Common user group. Users added to this user group will not have any additional permission. This type of user is available only in clusters with Kerberos authentication enabled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Default User Group	Description
presto	Common user group. Users added to this user group will not have any additional permission. This type of user is available only in clusters with Kerberos authentication enabled.
flume	Common user group. Users added to this user group will not have any additional permission. This type of user is available only in clusters with Kerberos authentication enabled.
launcher-job	MRS internal group, which is used to submit jobs using V2 APIs. This type of user is available only in clusters with Kerberos authentication enabled.

OS User Group	Description
wheel	Primary group of MRS internal running user **omm**.
ficommon	MRS cluster common group that corresponds to **compcommon** for accessing public resource files stored in the OS of the cluster.

- **Database user**

  MRS cluster system database users include OMS database users and DBService database users.

  WARNING

  Do not delete database users. Deleting them may cause the cluster or components to malfunction.

User Type	Default User	Initial Password	Description
OMS database	ommdba	dbChangeMe@123456	OMS database administrator who performs maintenance operations, such as creating, starting, and stopping.
	omm	ChangeMe@123456	User for accessing OMS database data

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

User Type	Default User	Initial Password	Description
DBService database	omm	dbserverAdmin@123	Administrator of the GaussDB database in the DBService component
	hive	HiveUser@	User for Hive to connect to the DBService database.
	hue	HueUser@123	User for Hue to connect to the DBService database.
	ranger	RangerUser@	User for Ranger to connect to the DBService database. This type of user is available only in clusters with Kerberos authentication enabled.
	sqoop	SqoopUser@	User for Loader to connect to the DBService database.

8.10.3 Managing MRS Cluster Roles

The administrator can create and manage different roles in Manager and use them to authorize access to Manager and components based on various service scenarios.

The Manager of MRS 3.x and later versions supports a maximum of 5,000 roles (including system built-in roles but excluding roles automatically created by tenants).

The Manager of MRS 2.x and earlier versions supports a maximum of 1000 roles.

Creating a Role

For MRS 3.*x* and later versions:

Step 1 Log in to MRS Manager and choose **System** > **Permission** > **Role**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 On the displayed page, click **Create Role** and fill in **Role Name** and **Description**.

The role name consists of 3 to 50 characters, including digits, letters, and underscores (_). It cannot be the same as an existing role name in the system. The role name cannot start with **Manager**, **System**, or **default**. For example, the role name cannot be **Manager_test**.

Step 3 In the **Configure Resource Permission** area, click the cluster whose permissions are to be added and select service permissions for the role.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

When setting permissions for a component, enter a resource name in the search text box in the upper right corner and click the search icon to view the search result.

The search result contains only directories, but not subdirectories. Search by keyword supports fuzzy match and is case-insensitive.

- For components (except HDFS and YARN) for which Ranger authorization has been enabled, the permissions of non-default roles on MRS Manager cannot take effect. You need to configure Ranger policies to assign permissions to user groups.
- If the resource requests of HDFS and YARN are beyond the Ranger policies, the ACL rules of the components still take effect.
- A maximum of 1000 permissions can be set for a component at a time.

Step 4 Click **OK**.

If the role name added in **Step 2** is displayed in the role list, the role is successfully added.

Step 5 After the role is created, create a user and bind the role to the user, or bind the role to an existing user. The user then has the relevant service permissions. For details, see **Creating an MRS Cluster User**.

----End

For MRS 2.x and earlier:

Step 1 Log in to MRS Manager and choose **System** > **Manage Role**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 On the displayed page, click **Create Role** and fill in **Role Name** and **Description**.

Role Name is mandatory and contains 3 to 30 characters. Only digits, letters, and underscores (_) are allowed. **Description** is optional.

Step 3 In **Permission**, set role permission.

1. Click **Service Name** and select a name in **View Name**.
2. Select one or more permissions.

- The **Permission** parameter is optional.
- If you select **View Name** to set component permissions, you can enter a resource name in the **Search** box in the upper right corner and click . The search result is displayed.
- The search result contains only directories, but not subdirectories. Search by keyword supports fuzzy match and is case-insensitive. Results on the next page can be searched.

Step 4 Click **OK**. Return to **Manage Role**.

If the role name added in **2** is displayed in the role list, the role is successfully added.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 After the role is created, create a user and bind the role to the user, or bind the role to an existing user. The user then has the relevant service permissions. For details, see **Creating an MRS Cluster User**.

----End

Managing Roles

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Go to the role management page.

- For MRS 3.x and later versions, choose **System** > **Permission** > **Role**.
- For MRS 2.x and earlier versions, choose **System** > **Manage Role**.

Step 3 In the role list, perform the following operations on the role as you need:

- View role information: Expand the expandable panel of the role to be viewed to check basic and permission information about the role.
- To modify role information, locate the target role, and click **Modify** in the **Operation** column. Default system roles cannot be modified.
- To export role information, click **Export All** in **TXT** or **CSV** format. The exported role information includes the role name and description.

  Exporting role information is available in MRS 3.*x* or later only.
- To delete a role, locate the row that contains the role and click **Delete**. To delete multiple roles in batches, select the target roles and click **Delete** above the role list. A role bound to a user cannot be deleted. To delete such a role, disassociate the role from the user by modifying the user first.

----End

Example: Creating a Manager Role (MRS 3.x and Later)

Step 1 Choose **System** > **Permission** > **Role**.

Step 2 On the displayed page, click **Create Role** and fill in **Role Name** and **Description**.

Step 3 In the **Configure Resource Permission** area, click **Manager** and set permissions for the role.

Manager permissions:

- Cluster
  - **view** permission: permission to view information on the **Cluster** page and view alarms and events under **O&M** > **Alarm**.
  - **management** permission: permission for management on the **Cluster** and **O&M** pages.
- User
  - **view** permission: permission to view information on pages under **System** > **Permission**.
  - **management** permission: permission for management on pages under **System** > **Permission**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Audit

  management permission: permission for management on the **Audit** page.

- Tenant
  - **view** permission: permission to view information on the **Tenant Resources** page. (Available in MRS 3.5.0 and later)
  - **management** permission: permission for management on the **Tenant** page and permission to view alarms and events under **O&M** > **Alarm**.
- System
  - **view** permission: permission to view all pages except those under **Permission** of the **System** module. (Available in MRS 3.5.0 and later)
  - **management** permission: permission for management on all pages except those under **Permission** of the **System** module and permission to view alarms and events under **O&M** > **Alarm**.

Step 4 Click **OK**.

----End

8.10.4 Managing MRS Cluster User Groups

Administrators can create and manage different user groups based on service scenarios on Manager. A user group is bound to a role to obtain operation permissions. After a user is added to a user group, the user can obtain the operation permissions of the user group. A user group can be used to classify users and manage multiple users.

In MRS 3.x or later clusters, MRS Manager supports up to 5000 user groups, including built-in groups.

In MRS 2.x or earlier clusters, MRS Manager supports up to 100 user groups, including built-in groups.

Prerequisites

- You have learned service requirements and created roles required by service scenarios.
- You have logged in to Manager.

Creating a User Group

For MRS 3.*x* and later versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User Group**.

Step 3 Above the user group list, click **Create User Group**.

Step 4 Set **Group Name** and **Description**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The group name contains 1 to 64 characters, including case-insensitive letters, digits, underscores (_), hyphens (-), and spaces. It cannot be the same as an existing user group name in the system.

Step 5 In the **Role** area, click **Add** to select a role and add it.

- For components (except HDFS and YARN) with Ranger authorization enabled, the permissions of non-default roles on MRS Manager do not take effect. You need to configure Ranger policies to assign permissions to user groups.
- When HDFS or YARN resource requests go beyond the scope of Ranger policies, the access control list (ACL) rules defined within the components continue to take effect.

Step 6 In the **User** area, click **Add** to select a user and add it.

Step 7 Click **OK**.

The user group is created.

----End

For MRS 2.x and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User Group**.

Step 4 Above the user group list, click **Create User Group**.

Step 5 Set **Group Name** and **Description**.

Group Name is mandatory and contains 3 to 20 characters. Only digits, letters, and underscores (_) are allowed. **Description** is optional.

Step 6 In **Role**, click **Select and Add Role** to select and add specified roles.

If you do not add the roles, the user group you are creating now does not have the permission to use MRS clusters.

Step 7 Click **OK**.

----End

Managing User Groups

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Go to the user group management page.

- For MRS 3.x and later versions, choose **System** > **Permission** > **User Group**.
- For MRS 2.x and earlier versions, Click **System**. In the **Permission** area, click **Manage User Group**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 In the user group list, perform the following operations on a user group as you need.

- Viewing user group information: By default, all user groups are displayed in the user group list. You can click the arrow on the left of a user group name to view details about the user group, including the user quantity, specific users, and bound roles of the user group.
- Modifying user group information: Locate the row that contains the user group to be modified and click **Modify**.
- Exporting user group information: Click **Export All** to export all user group information at a time in **TXT** or **CSV** format.

  Exporting user group information is supported only in MRS 3.*x* or later.

  The user group information contains the user group name, description, user list, and role list.
- Deleting a user group: Locate the row that contains the target user group and click **Delete**. To delete multiple user groups in batches, select the target user groups and click **Delete** above the user group list. A user group that contains users cannot be deleted. To delete such a user group, delete all its users by modifying the user group first.

----End

8.10.5 Managing MRS Cluster Users

8.10.5.1 Creating an MRS Cluster User

By default, only user **admin** has the highest operation permissions of Manager. Administrators need to create users on Manager and assign operation permissions to the users based on service requirements.

The FusionInsight Manager of an MRS 3.x or later cluster supports a maximum of 50000 users (including built-in users).

The MRS Manager of an MRS 2.x or earlier cluster supports a maximum of 1000 users.

Creating a User (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 On the **User** page, click **Create**.

Step 4 Set **Username**. The username can contain digits, letters, underscores (_), hyphens (-), and spaces. It is case-insensitive and cannot be the same as any existing username in the system or OS.

If the username contains special characters, such as periods (.) and at signs (@), the IAM user may fail to be synchronized.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Set **User Type** to **Human-Machine** or **Machine-Machine**.

- **Human-Machine** user: used for FusionInsight Manager O&M and component client operations. If you select this option, you also need to set **Password** and **Confirm Password**.
- **Machine-Machine** user: used for component application development. If you select this option, the password is randomly generated.

Step 6 In the **User Group** area, click **Add** to add one or more user groups to the list.

- If the selected user group has been bound to a role or a permission policy has been configured in Ranger, the user can obtain the corresponding permissions.
- After FusionInsight Manager is installed, some user groups generated by default have special permissions. Select desired user groups based on the descriptions on the UI.
- If existing user groups cannot meet your requirements, click **Create User Group** to create a user group. For details, see **Creating a User Group**.

Step 7 Select a group from the **Primary Group** drop-down list to create directories and files.

The drop-down list contains all groups selected in **User Group**.

A user can belong to multiple groups (including one primary group and multiple secondary groups). The primary group is set to facilitate maintenance and comply with the permission mechanism of the Hadoop community. The primary group has the same permission control functionality as other groups.

Step 8 In the **Role** area, click **Add** to bind roles to the user.

- Adding a role when you create a user can specify the user permissions.
- If the permissions granted to the user from the user group cannot meet service requirements, you can bind other created roles to the user. You can also click **Create Role** to create a role first. For details, see **Creating a Role**.

  It takes 3 minutes to make role permission assignment to the user take effect. If the permissions obtained from the user group are enough, you do not need to add a role.
- After Ranger authentication is enabled for a component, you need to configure Ranger policies to assign permissions to the user except the permissions of default user group or role.
- If a user is not added to a user group or assigned a role, the user cannot view information or perform operations after logging in to FusionInsight Manager.

Step 9 Enter information in **Description**.

Step 10 Click **OK**.

After a human-machine user is created, you need to change the initial password as prompted after logging in to FusionInsight Manager.

----End

Creating a User (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User**.

Step 4 On the **User** page, click **Create**.

Step 5 Configure parameters as prompted and enter a username in **Username**.

- A username that differs only in alphabetic case from an existing username is not allowed. For example, if **User1** has been created, you cannot create **user1**.
- When you use the user you created, enter the exactly correct username, which is case-sensitive.
- **Username** is mandatory and contains 3 to 20 characters. Only digits, letters, and underscores (_) are allowed.
- If the username contains special characters, such as periods (.) and at signs (@), the IAM user may fail to be synchronized.
- **root**, **omm**, and **ommdba** are reserved system user. Select another username.

Step 6 Set **User Type** to **Human-Machine** or **Machine-Machine**.

- **Human-machine** user: used for MRS Manager O&M scenarios and component client operation scenarios. If you select this user type, you need to enter a password and confirm the password in **Password** and **Confirm Password** accordingly.
- **Machine-machine** users: used for MRS application development scenarios. If you select this user type, you do not need to enter a password, because the password is randomly generated.

Step 7 In **User Group**, click **Select and Join User Group** to select user groups and add users to them.

- If roles have been added to user groups, the users can be granted with permissions of the roles.
- If you want to grant new users with Hive permissions, add the users to the Hive group.
- If a user needs to manage tenant resources, the user group must be assigned the **Manager_tenant** role and the role corresponding to the tenant.

Step 8 In **Primary Group**, select a group as the primary group for users to create directories and files. The drop-down list contains all groups selected in **User Group**.

Step 9 In **Assign Rights by Role**, click **Select and Add Role** to add roles for users based on onsite service requirements.

- If the permissions granted to the user from the user group cannot meet service requirements, you can bind other created roles to the user. It takes 3 minutes to make role permissions granted to the new user take effect.
- Adding a role when you create a user can specify the user permissions.
- A new user can access web UIs of HDFS, HBase, YARN, Spark, and Hue even when roles are not assigned to the user.

Step 10 In **Description**, provide description based on onsite service requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Description is optional.

Step 11 Click **OK**.

If a new user is used in the MRS cluster for the first time, for example, used for logging in to MRS Manager or using the cluster client, the password must be changed.

----End

8.10.5.2 Modifying MRS Cluster User Information

You can modify user information on Manager, including the user group, primary group, role permission assignment, and user description.

Modifying User Information (MRS 3.x or Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 Locate the row that contains the target user and click **Modify** in the **Operation** column.

Modify the parameters based on service requirements.

It takes three minutes at most for changes to user groups or role permissions to take effect.

Step 4 Click **OK**.

----End

Modifying User Information (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User**.

Step 4 In the row of a user to be modified, click **Modify**.

It takes three minutes at most for changes to user groups or assigned role permissions to take effect.

Step 5 Click **OK**. The modification is complete.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.5.3 Locking an MRS Cluster User

A user may be suspended for a long period of time due to service changes. For security purposes, you can lock such a user. You can lock a user in using either of the following methods:

- Automatic locking: You can set **Password Retries** in the password policy to automatically lock the user whose login attempts exceed this parameter value. For details, see **Configuring Password Policies for MRS Cluster Users**.
- Manual locking: You manually lock a user.

This section describes how to lock a user manually. Machine-machine users cannot be locked. For details about how to unlock an MRS cluster user, see **Unlocking a User Created on Manager**.

Impact on the System

A locked user cannot log in to Manager or perform identity authentication in the cluster. A locked user can be used only after being manually unlocked or the lock time expires.

Locking a User

For MRS 3.*x* and later versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 Locate the row that contains the target user and click **Lock** in the **Operation** column.

Step 4 In the window that is displayed, select **I have read the information and understand the impact**. Click **OK**.

----End

For MRS 2.x and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User**.

Step 4 In the row of a user you want to lock, click **Lock User**.

Step 5 In the window that is displayed, click **OK** to lock the user.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.5.4 Deleting an MRS Cluster User

Based on service requirements, you can delete system users that are no longer used on Manager.

- After a user is deleted, their provisioned TGT remains valid for up to 24 hours and can still be used for security authentication and system access.
- If a new user is created with the same name as a deleted user, the new user inherits all owner permissions of the deleted user. You may decide whether to delete the resources owned by a deleted user, such as HDFS files, based on service requirements.
- The default user **admin** cannot be deleted.

Deleting a Cluster User (MRS 3.x or Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 Locate the row that contains the target user, click **More**, and select **Delete**.

To delete users in batches, select multiple users and click **Delete**.

Step 4 In the displayed dialog box, click **OK**.

----End

Deleting a Cluster User (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User**.

Step 4 In the row that contains the user to be deleted, choose **More** > **Delete**.

Step 5 Click **OK**.

NOTE

If you want to create a new user with the same name as user A after deleting user A who has submitted a job on the client or MRS console, you need to delete user A's residual folders when deleting user A. Otherwise, the newly created user A may fail to submit a job.

To delete residual folders, log in to each Core node in the MRS cluster and run the following commands. In the following commands, **$user** indicates the folder named after the username.

```
cd /srv/BigData/hadoop/data1/nm/localdir/usercache/
rm -rf $user
```

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.5.5 Initializing MRS Cluster User Passwords

To reset a forgotten password or periodically change a public account password, administrator can initialize a new password on the Manager. The system user must then change the password upon their first login.

Impact on the System

If you have downloaded a user authentication file, download it again and obtain the keytab file after initializing the password of the MRS cluster user.

Initializing the Password of a Human-Machine User

For MRS 3.*x* and later versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 Locate the row that contains the target user, click **More**, and select **Initialize Password**. In the displayed dialog box, enter the password of the current login user and click **OK**. In the **Initialize Password** dialog box, click **OK**.

Step 4 Set **New Password** and **Confirm Password**, and click **OK**.

The password must meet the following complexity requirements:

- Contains at least 8 characters.
- Contains at least four types of the following: uppercase letters, lowercase letters, numbers, spaces, and special characters (`~!@#$%^&*()-_=+|[{}];',<.>/\?).
- Cannot be the same as the username or the username spelled backwards.
- Cannot be a common easily-cracked password.
- Cannot be the same as the password used in the latest *N* times. *N* indicates the value of **Repetition Rule** configured in **Configuring Password Policies for MRS Cluster Users**.

----End

For MRS 2.x and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User**.

Step 4 Locate the row that contains the user whose password is to be initialized, choose **More** > **Initialize password**, and change the password as prompted.

In the window that is displayed, enter the password of the current administrator account and click **OK**. Then in **Initialize password**, click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For the cluster, the default password complexity requirements are as follows:

- The password must contain 8 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
- The password cannot be the username or the username spelled backwards.

----End

Initializing the Password of a Machine-Machine User

Step 1 Prepare a client based on service conditions and log in to the node with the client installed.

Step 2 Run the following command to switch the user:

sudo su - omm

Step 3 Run the following command to switch to the client directory, for example, **/opt/client**:

cd /opt/client

Step 4 Run the following command to set environment variables:

source bigdata_env

Step 5 Run the following command to log in to the console as user **kadmin/admin**:

kadmin -p kadmin/admin

The default password of user **kadmin/admin** is **KAdmin@123**, which will expire upon your first login. Change the password as prompted and keep the new password secure.

Step 6 Run the following command to reset the password of a component running user. This operation takes effect on all servers:

cpw *Component running user name*

For example, **cpw oms/manager**.

For the cluster, the default password complexity requirements are as follows:

- The password must contain 8 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
- The password cannot be the username or the username spelled backwards.

----End

8.10.5.6 Downloading MRS Cluster User Credentials

If you develop big data applications and run them in an MRS cluster that requires Kerberos authentication, you need to prepare a user authentication file for accessing the MRS cluster. The keytab file in the authentication files is used for user authentication.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This topic describes how to download user authentication files and export the keytab file on MRS Manager.

Notes and Constraints

After a user password is changed, the exported keytab file becomes invalid. You must export a new keytab file.

Prerequisites

Before downloading the keytab file of a Human-Machine user, the password of the user must be changed at least once on the Manager portal or a client; otherwise, the downloaded keytab file cannot be used For details, see **Changing the Passwords for Manager Users of an MRS Cluster**.

Downloading Authentication Credential Files (for Versions Earlier Than MRS 3.5.0)

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

For MRS 3.*x* and later versions, choose **System** > **Permission** > **User**.

For versions earlier than MRS 3.*x*, choose **System** > **Permission** > **Manage User**.

Step 2 Locate the row that contains the user whose keytab file needs to be exported, choose **More** > **Download Authentication Credential**, specify the save path after the file is automatically generated, and keep the file properly.

The authentication credential includes the **krb5.conf** file of the Kerberos service.

After the authentication credential file is decompressed, you can obtain the following two files:

- The **krb5.conf** file contains the authentication service connection information.
- The **user.keytab** file contains user authentication information.

----End

Downloading Authentication Credential Files (for MRS 3.5.0 and Later)

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 Locate the row that contains the target user, and choose **More** > **Download Authentication Credential**.

Step 4 Select a location for downloading the authentication credential and set related parameters.

If the credential is downloaded to the server or a remote node, delete it after using it to prevent leakage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Browser**: Download the file to the local computer.
- **Server**: Download the file to the active OMS node of the cluster.

  The generated file is stored in the **/tmp/FusionInsight-Keytab/** directory on the active OMS node by default. If the path does not exist, it will be created. If the path already has an authentication credential file, it will be overwritten. For user **omm**, write permission for the path is required.

  After the file is generated, copy the downloaded package to another directory as the **omm** user.
- **Remote Node:** Download the file to a node other than the active OMS node. If you select this option, you need to set the following parameters:

  Table 8-63 Parameters for downloading to a remote node

Parameter	Example Value	Description
Save to Path	/tmp/FusionInsight-Keytab-Remote/	Path for storing the authentication credential file. If there is already a credential file in the path, it will be overwritten. The user for logging in to the remote node must have the write permission for the path.
Host IP Address	x.x.x.x	IP address of the remote node.
Host Port	22	Host port of the remote node.
Username	xxx	Username for logging in to the remote node. The user must have the write permission for the path.
Authentication Method	Password	You can choose one of the following methods: – **Password**: Use the password for login. – **None**: To use this method, passwordless login needs to be enabled.
Password	xxx	Enter the password when **Authentication Method** is set to **Password**. Password for logging in to the remote node.

Step 5 Click **OK**. Properly keep the file.

The authentication credential contains the **krb5.conf** file of Kerberos.

After the authentication credential file is decompressed, you can obtain the following two files:

- The **krb5.conf** file contains the authentication service connection information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The **user.keytab** file contains user authentication information.

----End

8.10.6 Unlocking an MRS Cluster User

8.10.6.1 Unlocking an LDAP User in the MRS Cluster

To unlock a user who has been locked out due to excessive incorrect password attempts, log in to the MRS cluster node and run commands to unlock the user, or use Manager to unlock the user if they were created on Manager.

If the service is abnormal, the internal user of the system may be locked. Unlock the user promptly, or the cluster cannot run properly. For the internal user list, see **MRS Cluster User Accounts**. Internal users cannot be unlocked on Manager.

Unlocking a User Created on Manager

For MRS 3.*x* and later versions:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User**.

Step 3 Locate the row that contains the target user and click **Unlock** in the **Operation** column.

Step 4 In the window that is displayed, select **I have read the information and understand the impact**. Click **OK**.

----End

For MRS 2.*x* and earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In the **Permission** area, click **Manage User**.

Step 4 In the row of a user to be unlocked, click **Unlock User**.

Step 5 In the window that is displayed, click **OK** to unlock the user.

----End

Unlocking an Internal User

This function is available only in MRS 3.*x* or later.

Step 1 Use the following method to confirm whether the internal system username is locked:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Method for obtaining the OLdap port number:
   a. Log in to Manager, choose **System** > **OMS** > **oldap** > **Modify Configuration**.
   b. The **LDAP Listening Port** parameter value is **oldap port**.
2. Method for obtaining the domain name:
   a. Log in to Manager, choose **System** > **Permission** > **Domain and Mutual Trust**.
   b. The **Local Domain** parameter value is the domain name.

      For example, the domain name of the current system is **9427068F-6EFA-4833-B43E-60CB641E5B6C.COM**.
3. Run the following command on each node in the cluster as user **omm** to query the number of password authentication failures:

   ldapsearch -H ldaps://*OMS Floating IP Address***:***OLdap port* **-LLL -x -D cn=root,dc=hadoop,dc=com -b krbPrincipalName=***Internal system username***@***Domain name***,cn=***Domain name***,cn=krbcontainer,dc=hadoop,dc=com -w** *Password of LDAP administrator* **-e ppolicy | grep krbLoginFailedCount**

   – To obtain the floating IP address of OMS, log in to the Master2 node remotely, and run the **ifconfig** command. In the command output, **eth0:wsom** indicates the floating IP address of OMS. Record the value of **inet**. If the floating IP address of OMS cannot be queried on the Master2 node, switch to the Master1 node to query and record the floating IP address. If there is only one Master node, query and record the IP address on the Master node.
   – Obtain the default password of LDAP administrator **cn=root,dc=hadoop,dc=com** by referring to **MRS Cluster User Accounts**.

   For example, run the following command to check the number of password authentication failures for user **oms/manager**:

   ldapsearch -H ldaps://10.5.146.118:21750 -LLL -x -D cn=root,dc=hadoop,dc=com -b krbPrincipalName=oms/manager@9427068F-6EFA-4833-B43E-60CB641E5B6C.COM,cn=9427068F-6EFA-4833-B43E-60CB641E5B6C.COM,cn=krbcontainer,dc=hadoop,dc=com -w *Password of user cn=root,dc=hadoop,dc=com* **-e ppolicy | grep krbLoginFailedCount**

   ```
   krbLoginFailedCount: 5
   ```
4. Log in to Manager, choose **System** > **Permission** > **Security Policy** > **Password Policy**.
5. Check the value of the **Password Retries** parameter. If the value is less than or equal to the value of **krbLoginFailedCount**, the user is locked.

   You can also check whether internal system users are locked by viewing run logs.

Step 2 Log in to the active management node as user **omm** and run the following command to unlock the user:

sh ${BIGDATA_HOME}/om-server/om/share/om/acs/config/unlockuser.sh --userName *Internal system username***--*****pwdPolicyName*****Password policy**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Example: **sh ${BIGDATA_HOME}/om-server/om/share/om/acs/config/unlockuser.sh --userName oms/manager**

Parameter **--pwdPolicyName** specifies the user password policy. This parameter is optional. When provided, it resets the user's password policy to the specified value and unlocks the user account.

If this script is used to unlock a human-machine user locked on MRS Manager and the user's password policy differs from the default, the **--pwdPolicyName** parameter must be provided to specify the correct policy. If the parameter is not specified, the password policy is automatically reset to the default. In other scenarios, you do not need to specify the **--pwdPolicyName** parameter

----End

8.10.6.2 Unlocking the LDAP Management Account of the MRS Cluster

If the LDAP user **cn=pg_search_dn,ou=Users,dc=hadoop,dc=com** and LDAP management accounts **cn=krbkdc,ou=Users,dc=hadoop,dc=com** and **cn=krbadmin,ou=Users,dc=hadoop,dc=com** are locked, the administrator must unlock these accounts.

Notes and Constraints

- If you enter an incorrect password for the LDAP user or management account for five consecutive times, the LDAP user or management account is automatically locked. Locked accounts are automatically unlocked after five minutes.
- This function is available only in MRS 3.*x* or later.

Procedure

Step 1 Log in to the active management node as user **omm**.

Step 2 Run the following command to go to the related directory:

cd ${BIGDATA_HOME}/om-server/om/ldapserver/ldapserver/local/script

Step 3 Run the following command to unlock the LDAP user or management account:

./ldapserver_unlockUsers.sh *USER_NAME*

In the command, *USER_NAME* indicates the name of the user to be unlocked.

For example, to unlock the LDAP management **account cn=krbkdc,ou=Users,dc=hadoop,dc=com**, run the following command:

./ldapserver_unlockUsers.sh krbkdc

After the script is executed, enter the password of user **krbkdc** next to **ROOT_DN_PASSWORD**. If the following information is displayed, the unlocking is successful:

```
Unlock user krbkdc successfully.
```

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.10.7 Configuring Password Policies for MRS Cluster Users

To keep up with service security requirements, you can set password security rules, user login security rules, and user locking rules on Manager.

Notes and Constraints

- This function is available only in MRS 3.1.2 or later.
- Password policies should be modified in alignment with service security requirements, since they directly impact user management security. Failure to do so may result in potential security risks.
- When a password policy is modified, the updated policy is enforced only after the user's password is changed. Therefore, the user must change their password for the new policy to take effect.
- This password policy applies to human-machine users created on MRS Manager. Up to 32 password policies can be created.

Adding a Password Policy

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **Security Policy** > **Password Policy**.

Step 3 Click **Add Password Policy** and modify the password policy as prompted.

For details about the parameters, see **Table 8-64**.

Table 8-64 Password policy parameters

Parameter	Example Value	Description
Password Policy Name	test	The value is a string of 3 to 32 characters, including case-insensitive letters, digits, underscores (_), and hyphens (-). It cannot start with a hyphen (-).
Minimum Password Length	8	Indicates the minimum number of characters a password contains. The value ranges from **8** to **32**.
Character Types	4	Indicates how many character types in the following types a password can contain at least: uppercase letters, lowercase letters, digits, spaces, and special characters (~`!?,.:;-_'(){}[]/<>@#$%^&*+|\=). The value can be **4** or **5**. The default value is **4**, which means that a password can contain uppercase letters, lowercase letters, digits, and special characters. If you set the parameter to **5**, a password can contain all the five character types mentioned above.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Password Retries	5	Indicates the number of consecutive wrong password attempts allowed before the system locks the user. The value ranges from **3** to **30**.
User Lock Duration (Min)	5	Indicates the time period in which a user is locked when the user lockout conditions are met. The value ranges from **5** to **120**.
Password Validity Period (Day)	90	Indicates the validity period of a password. The value ranges from **0** to **90**. **0** indicates that the password is permanently valid.
Repetition Rule	1	Indicates the number of previous passwords that cannot be reused when you change the password. The value ranges from **1** to **5**. The default value is **1**. This policy applies to only human-machine accounts.
Password Expiration Notification (Days)	5	Indicates the number of days in advance users are notified that their passwords are about to expire. After the value is set, if the difference between the cluster time and the password expiration time is smaller than this value, the user receives password expiration notifications. When a user logs in to Manager, a message is displayed, indicating that the password is about to expire and asking the user whether to change the password. The value ranges from **0** to *X* (*X* must be set to the half of the password validity period and rounded down). Value **0** indicates that no notification is sent. The default value is **5**.
Interval for Deleting Authentication Failure Records (Min)	5	Indicates the interval of retaining incorrect password attempts. The value ranges from **0** to **1440**. **0** indicates that incorrect password attempts are permanently retained, and **1440** indicates that incorrect password attempts are retained for one day.

Step 4 Click **OK** to save the settings.

A new user uses the default password policy. After a new password policy is created, you can manually select the password policy when creating a user. You

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

can modify the password policy of an existing user. For details, see **Modifying MRS Cluster User Information**.

Step 5 To delete a manually added password policy, perform the following operations:

Click **Delete** in the row that contains the target password policy. In the dialog box that is displayed, click **OK**.

The default password policy and the password policy that has been bound to a user cannot be deleted.

----End

Modifying a Password Policy

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Enter the password policy configuration page.

- For MRS 2.x and earlier versions: Choose **System** > **Password Policy Configuration**.
- For MRS 3.x and later versions, choose **System** > **Permission** > **Security Policy** > **Password Policy**, and click **Modify** in the row that contains the password policy you want to modify.

Step 3 Modify password policies as prompted.

Table 8-65 Password policy parameters

Parameter	Example Value	Description
Minimum Password Length	8	Indicates the minimum number of characters a password contains. The value ranges from 8 to 32. The default value is **8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Character Types	4	Indicates how many character types in the following types a password can contain at least: uppercase letters, lowercase letters, digits, spaces, and special characters (~`!?,.:;-_'(){}[]/<>@#$%^&*+|\=). • For MRS 3.x and later versions, the value can be **4** or **5**. The default value is **4**, which means that a password can contain uppercase letters, lowercase letters, digits, and special characters. If you set the parameter to **5**, a password can contain all the five character types mentioned above. • For MRS2.x and earlier versions, the value can be **3** or **4**. The default value **3** indicates that the password must contain at least three types of the following characters: uppercase letters, lowercase letters, digits, special characters, and spaces.
Password Retries	5	Indicates the number of consecutive wrong password attempts allowed before the system locks the user. The value ranges from **3** to **30**. The default value is **5**.
User Lock Duration (Min)	5	Indicates the time period in which a user is locked when the user lockout conditions are met. The value ranges from **5** to **120**. The default value is **5**.
Password Validity Period (Day)	90	Indicates the validity period (days) of a password. The value ranges from 0 to 90. Value **0** means that the password is permanently valid. The default value is **90**.
Repetition Rule	1	Indicates the number of previous passwords that cannot be reused when you change the password. The value ranges from **1** to **5**. The default value is **1**. This parameter is required for clusters of MRS 3.x or later. This policy applies to only human-machine accounts.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Password Expiration Notification (Days)	5	Indicates the number of days in advance users are notified that their passwords are about to expire. After the value is set, if the difference between the cluster time and the password expiration time is smaller than this value, the user receives password expiration notifications. When a user logs in to Manager, a message is displayed, indicating that the password is about to expire and asking the user whether to change the password. The value ranges from **0** to *X* (*X* must be set to the half of the password validity period and rounded down). Value **0** indicates that no notification is sent. The default value is **5**.
Interval for Deleting Authentication Failure Records (Min)	5	Indicates the interval (minutes) of retaining incorrect password attempts. The value ranges from 0 to 1440. **0** indicates that incorrect password attempts are permanently retained, and **1440** indicates that incorrect password attempts are retained for one day. The default value is **5**.

Step 4 Click **OK** to save the settings. After modifying a password policy, change the user password for the new policy to take effect.

In MRS 3.1.2 and later, users (except user **admin**) cannot modify the password policies bound to them. When the password policy bound to a user is modified, the system compares the remaining password validity period with the validity period defined in the new policy. If the remaining validity period exceeds the new policy's validity period, the validity period is reset to the new policy value. If the remaining validity period is shorter, the validity period remains unchanged.

----End

8.10.8 Configuring the Private Attribute of MRS Cluster Users

The private attribute of a user indicates that the user's password can be managed only by the user. The administrator cannot initialize or reset the password of a private user. If you do not want the administrator to reset your password, you can enable the private attribute. When a service user's private attribute is enabled, the service user becomes a private user, and the administrator cannot initialize the password of the private user.

User **admin** or administrators who are bound to the Manager_administrator role can configure the independent attribute on Manager so that common users (all service users in the cluster) can set or cancel their own independent attributes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the independent attribute option is toggled on, service users need to log in to the system and set the independent attribute.

Notes and Constraints

- Administrators cannot set or cancel the independent attribute of a user.
- Administrators cannot obtain the authentication credentials of independent users.
- This section applies only to MRS 3.*x* or later.

Prerequisites

You have obtained the required administrator username and password.

Enabling or Disabling User Private Attributes

Step 1 Log in to MRS Manager as user **admin** or a user bound to the Manager_administrator role.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **Security Policy** > **Independent Configurations**.

Step 3 Toggle on or off **Independent Attribute**, enter the password as prompted, and click **OK**.

Step 4 After the identity is authenticated, wait until the OMS configuration is modified and click **Finish**.

After the independent attribute is disabled:

- A service user who has this attribute can cancel it from the drop-down list of the username in the upper right corner of the page. The user cannot set the independent attribute again once it is cancelled. After the attribute is cancelled, existing independent tables will retain the attribute. However, the user cannot create independent tables again.
- Users without this attribute cannot set or cancel the attribute.

----End

Configuring the Independent Attribute

Step 1 Log in to MRS Manager as a service user.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

WARNING

- Administrators cannot initialize the password of a user after the independent attribute is set for the user. If the user password is forgotten, the password cannot be retrieved.
- User **admin** cannot set the independent attribute.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Move the cursor to the username in the upper right corner of the page.

Step 3 Select **Set Independent** or **Cancel Independent**.

- If the independent attribute is toggled on and has been set for the service user, **Cancel Independent** is displayed.
- If the independent attribute is toggled on but has been cancelled for the service user, **Set Independent** is displayed.
- If the independent attribute is toggled off but has been set for the service user, **Cancel Independent** is displayed.
- If the independent attribute is toggled off and has been cancelled for the service user, no option related to the independent attribute is displayed.

Step 4 Enter the password as prompted and click **OK**.

Step 5 After the identity is authenticated, click **OK** in the dialog box.

----End

8.11 Managing MRS Cluster Metadata

8.11.1 MRS Cluster Metadata Storage in an External Data Source

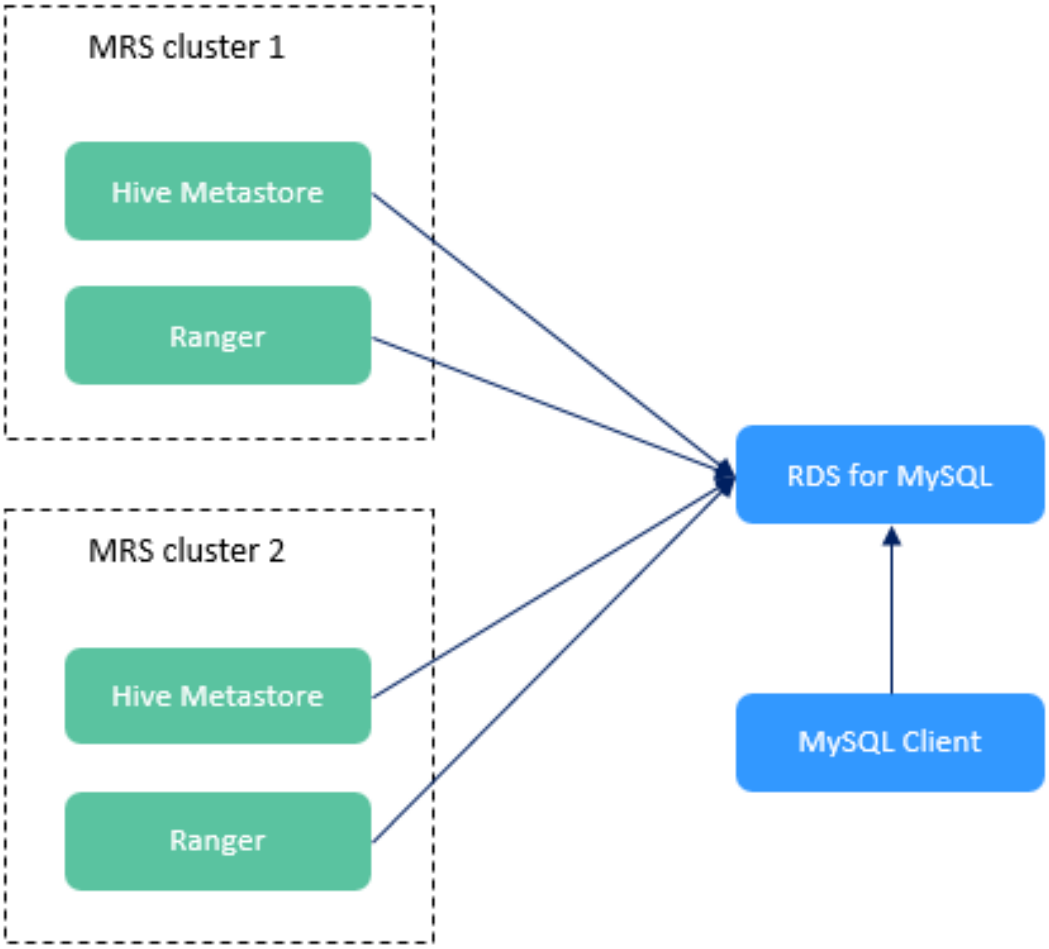
Metadata is data that describes other data, providing details such as its structure, storage location, and access permissions. In an MRS cluster, component metadata is stored by default within the local GaussDB database of the cluster. Deleting a cluster will also delete its metadata. To retain the metadata, you need to manually save it in advance.

MRS provides the data connection management function. This function allows metadata of components (such as Hive and Ranger) to be stored in external data sources, decoupling the data storage layer (such as HDFS) from compute engines (such as Spark and Flink).

For example, Hive metadata can be stored in an external relational database and will not be deleted when the current MRS cluster is deleted. In addition, multiple MRS clusters can share the same metadata.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 8-11 MRS cluster metadata storage in an external data source



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

External Data Connections Supported by MRS

Table 8-66 MRS external data connections

Data Connection Type	Description	Applicable Version	Supported Engine
RDS PostgreSQL database	RDS for PostgreSQL is designed for enterprise online transactional processing (OLTP) scenarios requiring complex SQL processing. It supports NoSQL data types (such as JSON, XML, and HStore) and geographic information system (GIS) data types, and is renowned for its reliability and data integrity. It is suitable for internet websites, location-based applications, and complex data object processing.	• MRS cluster version: clusters with the Hive component installed • PostgreSQL version: PostgreSQL 14	Hive
RDS MySQL database	RDS for MySQL is fully compatible with native MySQL, combining stability, reliability, and high performance. It features intelligent operations and maintenance, robust security, out-of-the-box usability, and automatic scaling.	• MRS cluster version: clusters with the Hive or Ranger component installed • MySQL versions: MySQL 5.7.*x* and MySQL 8.0	• Hive • Ranger

Notes and Constraints

- When Hive metadata is switched between different clusters, MRS synchronizes only the permissions in the metadata database of the Hive component. The permission model on MRS is maintained on MRS Manager. Therefore, when Hive metadata is switched between clusters, the permissions of users or user groups cannot be automatically synchronized to MRS Manager of another cluster.
- The VPC and subnet of the service for which an external data connection will be created must be the same as those of the MRS cluster to be interconnected.
- The RDS database instance interconnected with the MRS cluster cannot be deleted. Otherwise, the cluster will be abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.11.2 Storing Ranger Metadata to RDS

This topic describes how to switch the Ranger metadata of the existing cluster to the metadata stored in the RDS database. This operation enables MRS clusters to share the same metadata, retains the metadata when the cluster is deleted, and avoids Ranger metadata migration during cluster migration.

Notes and Constraints

When multiple MRS clusters connect to the same RDS database and share the same metadata, the clusters must be in the same mode.

Disabling Ranger Authentication for Cluster Components

This operation is required only for **MRS 3.1.0 or later**.

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > *Service name*.

Currently, the following components in an MRS 3.1.*x* or later cluster support Ranger authentication: HDFS, HBase, Hive, Spark, Impala, YARN, Kafka, and HetuEngine.

Step 3 In the upper right corner of the **Dashboard** page, click **More** and select **Disable Ranger**. If **Disable Ranger** is dimmed, Ranger authentication is disabled.

Step 4 (Optional) To use an existing authentication policy, perform this step to export the authentication policy on the Ranger web page. After the Ranger metadata is switched, you can import the existing authentication policy again. The following uses Hive as an example. After the export, a policy file in JSON format is generated in a local directory.

1. Log in to MRS Manager.
2. Choose **Cluster** > **Services** > **Ranger** to go to the Ranger service overview page.
3. Click **RangerAdmin** in the **Basic Information** area to go to the Ranger Web UI.

   For clusters with Kerberos authentication enabled, the **admin** user in Ranger belongs to the **User** type. To view all management pages, click the username in the upper right corner and select **Log Out** to log out of the system.

   Log in to the system as user **rangeradmin** (default password: **Rangeradmin@123**) or another user who has the Ranger administrator permissions.
4. Click the export button  in the row where the Hive component is located to export the authentication policy.
5. Click **Export**. After the export is complete, a policy file in JSON format is generated in a local directory.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Creating and Configuring an RDS DB Instance

Step 1 Log in to the RDS console and create an RDS DB instance. .

- To ensure network communications between the cluster and the MySQL database, create the instance in the same VPC and subnet as the MRS cluster.
- Security group rules of the RDS DB instance must allow inbound access from the MySQL database port (3306 by default).
- Ranger can interconnect with RDS for MySQL databases of the **MySQL 5.7.*x* and 8.0** versions only.

Step 2 In the navigation pane of the RDS console, choose **Instances**. Locate the row containing the RDS DB instance used by an MRS data connection, click **Log In** in the **Operation** column to log in to the DB instance as user **root**.

Step 3 On the home page of the instance, click **Create Database** to create a database.

WARNING

- If no new database is created, the MRS data connections will fail to set up.
- To connect MRS clusters in different modes to the same RDS database, you must create a new database.

Step 4 On the top of the page, choose **Account Management** > **User Management**.

- For clusters earlier than MRS 3.*x*, if the selected data connection type is **RDS MySQL database**, ensure that the database user is **root**. If the user is not **root**, create a user and grant permissions to the user by referring to **Step 5** to **Step 6**.
- For MRS 3.*x* or later clusters, if the selected data connection type is **RDS MySQL database**, the database user must not be **root**. In this case, create a user and grant permissions to the user by referring to **Step 5** to **Step 6**.

Step 5 Click **Create User** to create a non-**root** user, select all permissions in the **Global Permissions** area, and set other parameters as required.

If you are configuring an external RDS data connection for Ranger, you can select only the SELECT, INSERT, CREATE, RELOAD, CREATE USER, and GRANT permissions.

Step 6 On the top of the page, choose **SQL Operations** > **SQL Query**, switch to the target database by database name, and run the following SQL statements to grant permissions to the database user. In the following statements, *${db_name}* and *${db_user}* indicate the name of the database to be connected to MRS and the name of the new user, respectively.

```
grant all privileges on ${db_name}.* to '${db_user}'@'%' with grant option;
grant reload on *.* to '${db_user}'@'%' with grant option;
flush privileges;
```

----End

Creating an RDS Data Connection for an Existing MRS Cluster

Perform the following steps to create an RDS data connection for an existing MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to the MRS console.

Step 2 In the navigation pane, choose **Data Connections**.

Step 3 Click **Create Data Connection**.

Step 4 Set parameters by referring to **Table 8-67**.

Table 8-67 Parameters for creating a data connection

Parameter	Example Value	Description
Type	-	The type of an external source connection. Value options are as follows: • **RDS PostgreSQL database**. Clusters with Hive installed can connect to this type of database. • **RDS MySQL database**. Clusters with Hive or Ranger installed can connect to this type of database.
Name	newtest	The name of a data connection.
Database Instance	-	The RDS database instance. This instance must be created in RDS before being used here, and the database must have been created. For details, see **Creating and Configuring an RDS DB Instance**. Click **View DB Instance** to view the created DB instances.
Database	dataname	The name of the database to be connected.
Username	datauser	The username for logging in to the database to be connected.
Password	-	The password for logging in to the database to be connected.

If the selected data connection type is **RDS MySQL database**, ensure that the database user is **root**. If the user is not **root**, perform operations by referring to **Creating and Configuring an RDS DB Instance**.

Step 5 Click **OK**.

----End

Configuring a Ranger Data Connection

Step 1 Log in to the MRS console.

Step 2 Click the name of a cluster to go to the cluster details page.

Step 3 Click **Manage** on the right of **Data Connection** to go to the data connection configuration page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Click **Configure Data Connection** and set related parameters.

- **Type**: Ranger
- **Module Type**: Ranger metadata
- **Connection Type**: RDS MySQL database
- **Connection Instance**: Select a created RDS MySQL DB instance. For details about how to create a data connection, see **Creating an RDS Data Connection for an Existing MRS Cluster**.

Step 5 Select **I understand the consequences of performing the scale-in operation** and click **Test**.

Step 6 After the test is successful, click **OK** to complete the data connection configuration.

Step 7 Log in to MRS Manager. For details, see **Accessing MRS Manager**.

Step 8 Choose **Cluster** > **Services** > **Ranger** to go to the Ranger service overview page.

Step 9 Choose **More** > **Restart Service** or **More** > **Service Rolling Restart**.

If you choose **Restart Service**, services will be interrupted during the restart. If you select **Service Rolling Restart**, rolling restart can minimize the impact or do not affect service running.

Restarting Ranger will affect the permissions of all components controlled by Ranger and may affect service running. Restart Ranger when the cluster is idle or during off-peak hours. Before the Ranger component is restarted, the policies in the Ranger component still take effect.

Step 10 Enable Ranger authentication for the component to be authenticated. The Hive component is used as an example.

Currently, the following components in an MRS 3.1.*x* or later cluster support Ranger authentication: HDFS, HBase, Hive, Spark, Impala, Storm, Kafka, HetuEngine, and YARN.

1. Log in to MRS Manager and choose **Cluster** > **Services** > *Service Name*.
2. In the upper right corner of the **Dashboard** page, click **More** and select **Enable Ranger**.

Step 11 Log in to the Ranger Web UI as a user with the Ranger administrator permissions and click  for the Hive component.

Step 12 Import parameters.

- Click **Select file** and select the authentication policy file downloaded in **Step 4.5**.
- Select **Merge If Exist Policy**.

Step 13 Restart the component for which Ranger authentication is enabled.

1. Log in to MRS Manager. For details, see **Accessing MRS Manager**.
2. Choose **Cluster** > **Services** > *Service name*. The service overview page is displayed.
3. Choose **More** > **Restart Service** or **More** > **Service Rolling Restart**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If you choose **Restart Service**, services will be interrupted during the restart. If you select **Service Rolling Restart**, rolling restart can minimize the impact or do not affect service running.

----End

Verifying Ranger Metadata Storage in an External Database

Step 1 For a cluster with Kerberos authentication disabled, the Ranger data source can be accessed by Unix users of the cluster by default. To enable the data source to be accessed by the MRS Manager LDAP users, perform the following operations:

1. Log in to MRS Manager by referring to **Accessing MRS Manager**. Choose **Cluster** > **Services** > **Ranger**, click **Configurations** and then **All Configurations**, click **UserSync (Role)** and then **Customization**, and add the following custom parameters to set the Ranger data source to the MRS Manager LDAP:
   - For versions later than MRS 3.1.0, in the **ranger.usersync.config.expandor** area, set **ranger.usersync.sync.source** to **ldap** and **ranger.usersync.cookie.enabled** to **false**.
   - For MRS 3.1.0, in the **ranger.ugsync.site.customized.configs** area, set **ranger.usersync.sync.source** to **ldap** and **ranger.usersync.cookie.enabled** to **false**.
2. Click **Save**.
3. In the upper right corner of the Ranger **Dashboard** page, click **More** and choose **Synchronize Configuration**.
4. On the Ranger instance page, select the **UserSync** instance and choose **More** > **Restart Instance**.
5. On the Ranger **Dashboard** page, click **RangerAdmin** and choose **Settings** > **Users/Groups/Roles** to check whether LDAP users exist.

Step 2 Enable Ranger authentication for Hive.

On MRS Manager, choose **Cluster** > **Services** > **Hive**. In the upper right corner of the page, click **More** and check whether Ranger authentication is enabled for Hive. If Ranger authentication is not enabled, select **Enable Ranger** to enable it.

Step 3 On MRS Manager of the cluster, choose **System** > **Permission** > **User** and add a human-machine user, for example, user **test**. Users in a cluster with Kerberos authentication enabled must have the permissions to create Hive tables. That is, the user must belong to either the **hive** or **hadoop** user group.

Step 4 Log in to the node where the Hive client is installed and run the following commands:

Switch to the client installation directory.

```
cd Client installation directory
```

Load the environment variables.

```
source bigdata_env
```

Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
kinit Service user
```

Step 5 Run the following command to log in to the Hive client:

- For clusters with Kerberos authentication enabled:
  ```
  beeline
  ```
- For clusters with Kerberos authentication disabled:
  ```
  beeline -n Service user
  ```

Step 6 Run the following commands to create a Hive table and insert data to the table:

Create a table:

```
create table user_info(id string,name string,gender string,age int,addr string);
```

Insert data:

```
insert into table user_info(id,name,gender,age,addr) values("12005000201","A","man",19,"city");
```

Step 7 On MRS Manager of the cluster, choose **System** > **Permission** > **User** and add another human-machine user, for example, user **test1**. Users in a cluster with Kerberos authentication enabled can have only the permissions on the **hive** user group.

Step 8 Log in to the Ranger Web UI as user **rangeradmin** (for a cluster with Kerberos authentication enabled) or **admin** (for a cluster with Kerberos authentication disabled), and add the permission to query the table created in **Step 6** for the user.

1. On the Service Manager page, click **Hive** in **Hadoop SQL**, click **Add New Policy** to add a Hive permission control policy, set the following parameters, and click **Add**.
   - **Policy Name**: Enter a policy name.
   - **database**: Set it to **default**.
   - **table**: Set it to the name of the table to be accessed, for example, **user_info**.
   - **column**: Set it to *****.
   - In the **Allow Conditions** area, select the user you want to authorize, for example, **test1**, in the **Select User** column, and select the **select** permission in the **Permissions** column.
2. For a cluster with Kerberos authentication enabled, return to the Service Manager page, click **hacluster** in **HDFS**, click **Add New Policy** to add an HDFS permission control policy, set the following parameters, and click **Add**.
   - **Policy Name**: Enter a policy name.
   - **Resource Path**: Enter the specific path of the table to be accessed in HDFS, for example, **/user/hive/warehouse/user_info**.
   - In the **Allow Conditions** area, select the user you want to authorize, for example, **test1**, in the **Select User** column, and select **Read** and **Execute** in the **Permissions** column.

Step 9 Log in to the Hive client as the user created in **Step 7** and run the following command to query data in the table created in **Step 6**:

select * from *user_info*;

If the table data can be queried, the Ranger metadata is successfully stored to the external RDS MySQL database.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 8-12 Hive table data queried



----End

8.11.3 Storing Hive Metadata to RDS

This topic describes how to switch the Hive metadata of the existing cluster to the metadata stored in the RDS database. This operation enables MRS clusters to share the same metadata, retains the metadata when the cluster is deleted, and avoids Hive metadata migration during cluster migration.

Creating and Configuring an RDS DB Instance

Step 1 Log in to the RDS console and create an RDS DB instance. .

- To ensure network communications between the cluster and the MySQL or PostgreSQL database, create the instance in the same VPC and subnet as the cluster.
- Security group rules of the RDS DB instance must allow inbound access from MySQL (default port 3306) and PostgreSQL (default port 5432) database ports.

  For example, click the instance name on the RDS console to go to the instance management page. In the **Connection Information** area, click the name next to **Security Group**. On the page that is displayed, click the **Inbound Rules** tab, and click **Add Rule**. In the displayed **Add Inbound Rule** dialog box, in the **Protocol & Port** area, set the protocol to TCP and port to 3306. In the **Source** area, select **IP address** and enter the IP addresses of all nodes where the MetaStore instances of Hive are located.
- Hive can interconnect with RDS for MySQL and PostgreSQL databases. The supported versions are **MySQL 5.7.*x* and 8.0** and **PostgreSQL 14**.

Step 2 In the navigation pane of the RDS console, choose **Instances**. Locate the row containing the RDS DB instance used by an MRS data connection, click **Log In** in the **Operation** column to log in to the DB instance as user **root**.

Step 3 On the home page of the instance, click **Create Database** to create a database.

WARNING

If no new database is created, the MRS data connections will fail to configure.

Step 4 On the top of the page, choose **Account Management** > **User Management**.

- For clusters earlier than MRS 3.*x*, if the selected data connection type is **RDS MySQL database**, ensure that the database user is **root**. If the user is not **root**, create a user and grant permissions to the user by referring to **Step 5** to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 3.*x* or later clusters, if the selected data connection type is **RDS MySQL database**, the database user must not be **root**. In this case, create a user and grant permissions to the user by referring to **Step 5** to **Step 6**.

Step 5 Click **Create User** to create a non-**root** user, select all permissions in the **Global Permissions** area, and set other parameters as required.

Step 6 On the top of the page, choose **SQL Operations** > **SQL Query**, switch to the target database by database name, and run the following SQL statements to grant permissions to the database user. In the following statements, *${db_name}* and *${db_user}* indicate the name of the database to be connected to MRS and the name of the new user, respectively.

```
grant all privileges on ${db_name}.* to '${db_user}'@'%' with grant option;
grant reload on *.* to '${db_user}'@'%' with grant option;
flush privileges;
```

----End

Creating an RDS Data Connection for an Existing MRS Cluster

Perform the following steps to create an RDS data connection for an existing MRS cluster.

Step 1 Log in to the MRS console.

Step 2 In the navigation pane, choose **Data Connections**.

Step 3 Click **Create Data Connection**.

Step 4 Set parameters by referring to **Table 8-68**.

Table 8-68 Parameters for creating a data connection

Parameter	Example Value	Description
Type	-	The type of an external source connection. Value options are as follows: • **RDS PostgreSQL database**. Clusters with Hive installed can connect to this type of database. • **RDS MySQL database**. Clusters with Hive or Ranger installed can connect to this type of database.
Name	newtest	The name of a data connection.
Database Instance	-	The RDS database instance. This instance must be created in RDS before being used here, and the database must have been created. For details, see **Creating and Configuring an RDS DB Instance**. Click **View DB Instance** to view the created DB instances.
Database	dataname	The name of the database to be connected.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Username	datauser	The username for logging in to the database to be connected.
Password	-	The password for logging in to the database to be connected.

If the selected data connection type is **RDS MySQL database**, ensure that the database user is **root**. If the user is not **root**, perform operations by referring to **Creating and Configuring an RDS DB Instance**.

Step 5 Click **OK**.

----End

Configuring a Hive Data Connection

This function is not supported in MRS 3.0.5.

Step 1 Log in to the MRS console.

Step 2 In the navigation pane, choose **Active Clusters**.

Step 3 Click the name of a cluster to go to the cluster details page.

Step 4 On the **Dashboard** tab page, click **Manage** next to **Data Connection**.

Step 5 On the **Data Connections** page, the data connections of the cluster are displayed. You can click **Disassociate** to delete a data connection.

Step 6 If there is no associated data connection on the **Data Connection** dialog box, click **Configure Data Connection** to add a connection.

Only one data connection can be configured for a module type. For example, after a data connection is configured for Hive metadata, no other data connection can be configured for it. If no module type is available, the **Configure Data Connection** button is unavailable.

Table 8-69 Configuring a Hive data connection

Parameter	Description
Type	Hive
Module Type	Hive metadata
Data Connection Type	• RDS PostgreSQL database (supported for clusters of MRS 1.9.*x*) • RDS MySQL database • Local database

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Instance	This parameter is valid only when **Data Connection Type** is set to **RDS PostgreSQL database** or **RDS MySQL database**. Select the name of the connection between the MRS cluster and the RDS database. This instance must be created before being referenced here. You can click **Create Data Connection** to create a data connection. For details, see **Creating an RDS Data Connection for an Existing MRS Cluster**.

Step 7 Click **Test** to test connectivity of the data connection.

Step 8 After the data connection is successful, click **OK**.

Step 9 If the RDS for PostgreSQL database is interconnected, upload the open- source PostgreSQL driver package to all nodes where MetaStore instances are deployed to replace the existing driver package in the cluster.

Upload the driver package **postgresql-42.2.5.jar** to the **${BIGDATA_HOME}/third_lib/Hive** directory on all MetaStore instance nodes.

Step 10 On the cluster details page, click the **Components** tab and then **Hive**.

For MRS 3.2.0-LTS or later, if multiple MRS clusters are interconnected with the same RDS database, choose **Service Configuration**, change **Basic** to **All** in the search area, choose **MetaStore** > **Customization** in the navigation pane, add **metastore.whether.query.all.tokens** to the **hive.metastore.customized.configs** parameter, set the value to **false**, and click **Save Configuration**.

Step 11 On the **Service Status** tab page, choose **More** > **Restart Service** and click **OK** to restart the Hive service.

- If IAM users are not synchronized, click the **Dashboard** tab on the cluster details page, click **Synchronize** on the right of **IAM User Sync**, and then restart the Hive service.
- Hive will create necessary database tables in the specified database after restart. (If tables already exist, they will not be created again.)

----End

Verifying Hive Metadata Storage in an External Database

Step 1 Log in to the node where the Hive client is installed and run the following commands. The authenticated user must have the permission to create Hive tables.

Switch to the client installation directory.

```
cd Client installation directory
```

Load the environment variables.

```
source bigdata_env
```

Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
kinit Service user
```

Step 2 Run the following command to log in to the Hive client:

```
beeline
```

Step 3 Run the following commands to create a Hive table and insert data to the table:

Create a table:

```
create table user_info(id string,name string,gender string,age int,addr string);
```

Insert data:

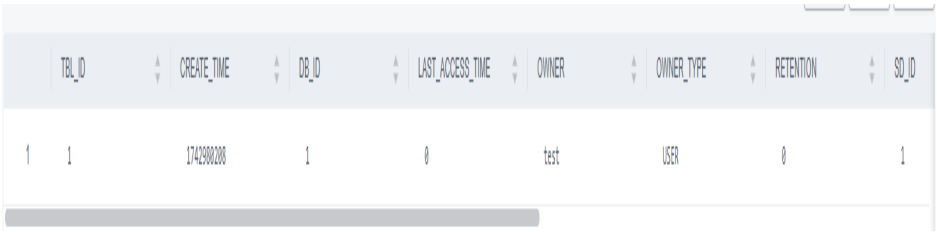
```
insert into table user_info(id,name,gender,age,addr) values("12005000201","A","man",19,"city");
```

Step 4 On the SQL query page of the database of the interconnected RDS for MySQL instance, run the following command:

```
select * from tbls;
```

If the query result contains information about the Hive table created in **Step 3**, the Hive metadata is successfully stored to the external RDS MySQL database.

Figure 8-13 Table information queried



----End

8.11.4 Managing MRS Cluster Data Connections

This topic describes how to create, view, and delete a cluster data connection on the MRS console.

Creating a Data Connection

Step 1 Log in to the MRS management console. In the navigation pane, choose **Data Connections**.

Step 2 Click **Create Data Connection**.

For details about how to configure an RDS for MySQL data connection, see **Creating an RDS Data Connection for an Existing MRS Cluster**.

Step 3 Click **OK**.

----End

Checking Data Connection Information

Step 1 Log in to the MRS management console. In the navigation pane, choose **Data Connections**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 In the data connection list, click the desired data connection. On the page that is displayed, view its details.

----End

Deleting a Data Connection

Step 1 Log in to the MRS management console. In the navigation pane, choose **Data Connections**.

Step 2 In the **Operation** column of the data connection list, click **Delete** in the row where the data connection to be deleted is located.

If the selected data connection has been associated with a cluster, the deletion does not affect the cluster.

----End

8.12 Managing Static Service Resources in an MRS Cluster

8.12.1 Overview of Static Service Resources

Introduction

A cluster allocates static service resources to services Flume, HBase, HDFS, Kafka (Kafka supports static service pools only in MRS 3.2.0 or later), and YARN. The total volume of computing resources allocated to each service is fixed, and they are static. A tenant can exclusively use or share a service to obtain the resources required for running this service.

Static Service Pool

Static service pools are used to specify service resource configurations.

Static service pools centrally manage resources that can be used by each service.

- Limits the total number of resources that can be used by each service. Specifically, the total number of CPU, I/O, and memory resources can be configured on the nodes where services Flume, HBase, HDFSKafka (Kafka supports static service pools only in MRS 3.2.0 or later), and YARN are deployed.
- Isolates the resources of services in a cluster from those of other services. In this way, the load of one service has very limited impact on other services.

Scheduling Mechanism

The time-based dynamic resource scheduling mechanism enables different volumes of static resources to be configured for services at different time, optimizing service running environments and improving the cluster efficiency.

In a complex cluster environment, multiple services share resources in the cluster, but the resource service period of each service may be different.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The following use a bank customer as an example:

- The HBase query service is heavy in the daytime.
- The query service is light, but the Hive analysis service is heavy at night.

If fixed resources are allocated to each service, the following problems may occur:

- The query service cannot obtain sufficient resources while the resources for the analysis service are idle in the daytime.
- The analysis service cannot obtain sufficient resources while the resources for the query service are idle at night.

As a result, the cluster resource utilization is low and the service capability is weak. Resolve the problem in the following ways:

- Sufficient resources need to be configured for HBase in the daytime.
- Sufficient resources need to be configured for Hive at night.

The time-based dynamic scheduling mechanism can efficiently utilize resources and run tasks.

8.12.2 Configuring Static Resources for an MRS Cluster

You can adjust resource base on Manager and customize resource configuration groups if you need to control service resources used on each node in a cluster or the available CPU or I/O quotas on each node at different time segments.

Impact on the System

- After a static service pool is configured, the configuration of affected services expires. You need to restart the services. Services are unavailable during restart.
- After a static service pool is configured, the maximum number of resources used by each service and role instance cannot exceed the upper limit.

Configuring Static Resources (MRS 3.x and Later)

Modify the Resource Adjustment Base

Step 1 Log in to the MRS Manager and choose **Cluster** > **Static Service Pool Configurations**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Click **Configuration** in the upper right corner. The page for configuring resource pools is displayed.

Step 3 Change the values of **CPU (%)** and **Memory (%)** in the **System Resource Adjustment Base** area.

Modifying the system resource adjustment base changes the maximum physical CPU and memory usage on nodes by services. If multiple services are deployed on the same node, the maximum physical resource usage of all services cannot exceed the adjusted CPU or memory usage.

Step 4 Click **Next**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

To modify parameters again, click **Previous**.

Modify the Default Resource Configuration Group

Step 5 Click **default**. In the **Configure weight** table, set **CPU LIMIT(%)**, **CPU SHARE(%)**, **I/O(%)**, and **Memory(%)** for each service.

- The sum of **CPU LIMIT(%)** and **CPU SHARE(%)** values configured for all services can exceed 100%.
- The sum of **I/O(%)** values configured for all services can exceed 100% but cannot be 0.
- The sum of **Memory(%)** values configured for all services can exceed, fall below, or equal 100%.
- **Memory(%)** settings cannot take effect dynamically and can only be modified in the default configuration group.
- **CPU LIMIT(%)** specifies the proportion of CPU cores available to a service relative to the total number of cores allocated to the associated nodes.
- **CPU SHARE(%)** specifies the proportion of CPU core time assigned to a service relative to the time allocated to other services sharing the same core. It represents the time ratio when multiple services contend for the same CPU core.

Step 6 Click **Generate detailed configurations based on weight configurations**. FusionInsight Manager generates the actual values of the parameters in the default weight configuration table based on the cluster hardware resources and allocation information.

Step 7 Click **OK**.

In the displayed dialog box, click **OK**.

Add a Customized Resource Configuration Group

Step 8 Determine whether to automatically adjust resource configurations at different time segments.

- If yes, go to **Step 9**.
- If no, use the default configurations, and no further action is required.

Step 9 Click **Configuration**, change the system resource adjustment base values, and click **Next**.

Step 10 Click **Add** to add a resource configuration group.

Step 11 In **Step 1: Scheduling Time**, click **Configuration**. The time policy configuration page is displayed.

Modify the following parameters based on service requirements and click **OK**.

- **Repeat**: If selected, the resource configuration group will run repeatedly based on the scheduling period. If not selected, set the date and time when the resource configuration group can be applied.
- **Repeat Policy**: You can select **Daily**, **Weekly**, or **Monthly**. This parameter is available only when **Repeat** is selected.
- **Between**: indicates the time period between the start time and end time when the resource configuration is applied. Set a unique time range. If the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

time range overlaps with that of an existing group of resource configuration, the time range cannot be saved.

Description

- The **default** resource configuration group takes effect in all undefined time segments.
- The newly added resource configuration group is a parameter set that takes effect dynamically in a specified time range.
- The newly added resource configuration group can be deleted. A maximum of four resource configuration groups that take effect dynamically can be added.
- Select a repetition policy. When the configured end time precedes the start time, the system automatically interprets the end time as occurring on the following day. For example, when a validity period is defined from 22:00 to 06:00, the resource configuration becomes effective at 22:00 on the current day and remains in effect until 06:00 on the following day.
- When multiple configuration groups are assigned different repeat policy types, their time ranges can overlap. The policy types are prioritized in ascending order as follows: daily, weekly, and monthly. For example, consider two resource configuration groups, one using a monthly policy and the other a daily policy. Their application time ranges overlap ([04:00–07:00] and [06:00–08:00]). In such cases, the configuration of the group with the monthly policy is applied.
- When multiple resource configuration groups are assigned the same repeat policy type, the time ranges defined for different dates can overlap. For example, consider two weekly resource configuration groups. You can configure them with identical time ranges on different days, such as 04:00–07:00 on Monday and 04:00–07:00 on Wednesday.

Step 12 Modify the resource configuration of each service in **Step 2: Weight Configuration**.

Step 13 Click **Generate detailed configuration**. FusionInsight Manager generates the actual values of the parameters in the default weight configuration table based on the cluster hardware resources and allocation information.

Step 14 Click **OK**.

In the displayed dialog box, click **OK**.

----End

Configuring Static Resources (MRS 2.x and Earlier)

Step 1 Modify the system resource adjustment base.

1. Log in to MRS Manager and click **System**. In the **Resource** area, click **Configure Static Service Pool**.

   For details about how to log in to MRS Manager, see **Accessing MRS Manager**.
2. Click **Configuration**. The service pool configuration group management page is displayed.
3. In the **System Resource Adjustment Base** area, change the values of **CPU(%)** and **Memory(%)**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Modifying **System Resource Adjustment Base** limits the maximum physical CPU and memory resource percentage of nodes that can be used by the Flume, HBase, HDFS, Impala and YARN services. If multiple services are deployed on the same node, the maximum physical resource usage of all services cannot exceed the adjusted CPU or memory usage.

4. Click **Next**.

   If you need to modify the parameters again, click **Previous** in the lower part of the page.

Step 2 Modify the **default** configuration group of the service pool.

1. In the **Service Pool Configuration** table, set **CPU LIMIT(%)**, **CPU SHARE(%)**, **I/O(%)**, and **Memory(%)** for the Flume, HBase, HDFS, Impala, and YARN services.

- The sum of **CPU LIMIT(%)** used by all services can exceed 100%.
- The sum of **CPU SHARE(%)** and **I/O(%)** used by all services must be 100%. For example, if CPU resources are allocated to the HDFS and YARN services, the total CPU resources allocated to the two services are 100%.
- The sum of **Memory(%)** used by all services can be greater than, smaller than, or equal to 100%.
- **Memory(%)** settings cannot take effect dynamically and can only be modified in the default configuration group.

2. Click in the blank area of the page to complete the editing. MRS Manager generates the correct values of service pool parameters in the **Detailed Configuration** area based on the cluster hardware resources and allocation information.
3. You can click the edit icon on the right of **Detailed Configuration** to modify the parameter values of the service pool based on service requirements.

   In the **Service Pool Configuration** area, click the specified service name. The **Detailed Configuration** area displays only the parameters of the service. Manually changing parameter values does not refresh the service resource usage. In added configuration groups, the configuration group numbers of the parameters that take effect dynamically will be displayed. For example, HBase: RegionServer: dynamic-config1.RES_CPUSET_PERCENTAGE. The parameter functions do not change.

Table 8-70 Parameters of the static service pool

Parameter	Description
– RES_CPUSET_PERCENTAGE – dynamic-configX.RES_CPUSET_PERCENTAGE	Configures the service CPU percentage.
– RES_CPU_SHARE – dynamic-configX.RES_CPU_SHARE	Configures the service CPU usage share.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
– RES_BLKIO_WEIGHT – dynamic-configX.RES_BLKIO_WEIGHT	Configures service I/O usage.
HBASE_HEAPSIZE	Configures the maximum JVM memory for RegionServer.
HADOOP_HEAPSIZE	Configures the maximum JVM memory of a DataNode.
yarn.nodemanager.resource.memory-mb	Configures the memory that can be used by NodeManager on the current node.
dfs.datanode.max.locked.memory	Configures the maximum memory that can be used by a DataNode as the HDFS cache.
FLUME_HEAPSIZE	Configures the maximum JVM memory that can be used by each Flume instance.
IMPALAD_MEM_LIMIT	Configures the maximum memory that can be used by an Impalad instance.

Step 3 Add a customized resource configuration group.

1. Determine whether to automatically adjust resource configurations based on the time.

   If yes, go to **Step 3.2**.

   If no, go to **Step 4**.

2. Click  to add a resource configuration group. In the **Scheduling Time** area, click the edit icon. The time policy configuration page is displayed.

   Modify the following parameters based on service requirements and click **OK**.

   – **Repeat**: If selected, the resource configuration group will run repeatedly based on the scheduling period. If not selected, set the date and time when the resource configuration group can be applied.

   – **Repeat Policy**: You can select **Daily**, **Weekly**, or **Monthly**. This parameter is available only when **Repeat** is selected.

   – **Between**: indicates the time period between the start time and end time when the resource configuration is applied. Set a unique time range. If the time range overlaps with that of an existing group of resource configuration, the time range cannot be saved. This parameter is available only when **Repeat** is selected.

   Description

   – The **default** resource configuration group takes effect in all undefined time segments.

   – The newly added resource configuration group is a parameter set that takes effect dynamically in a specified time range.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The newly added resource configuration group can be deleted. A maximum of four resource configuration groups that take effect dynamically can be added.
- Select a repetition policy. When the configured end time precedes the start time, the system automatically interprets the end time as occurring on the following day. For example, when a validity period is defined from 22:00 to 06:00, the resource configuration becomes effective at 22:00 on the current day and remains in effect until 06:00 on the following day.
- When multiple configuration groups are assigned different repeat policy types, their time ranges can overlap. The policy types are prioritized in ascending order as follows: daily, weekly, and monthly. For example, consider two resource configuration groups, one using a monthly policy and the other a daily policy. Their application time ranges overlap ([04:00–07:00] and [06:00–08:00]). In such cases, the configuration of the group with the monthly policy is applied.
- When multiple resource configuration groups are assigned the same repeat policy type, the time ranges defined for different dates can overlap. For example, consider two weekly resource configuration groups. You can configure them with identical time ranges on different days, such as 04:00–07:00 on Monday and 04:00–07:00 on Wednesday.

3. On the **Service Pool Configuration** page, modify the resource configuration of each service. Click the blank area on the page to complete the editing, and go to **Step 4**.

   You can click the edit icon on the right of **Service Pool Configuration** to modify the parameters. Click the edit icon in the **Detailed Configuration** area to manually update the parameter values generated by the system based on service requirements.

Step 4 Saves the settings.

Click **Save**. In the **Save Configuration** dialog box, select **Restart the affected services or instances**. Click **OK** to save the settings and restart related services.

After the system displays "Operation succeeded", click **Finish**. The service is started successfully.

----End

8.12.3 Checking the Static Resources of an MRS Cluster

The big data management platform can manage and isolate service resources that are not running on YARN using static service resource pools. The system supports time-based automatic adjustment of static service resource pools. This enables the cluster to automatically adjust the parameter values at different periods to ensure more efficient resource utilization.

System administrators can view the monitoring indicators of resources used by each service in the static service pool on Manager. The monitoring indicators are as follows:

- CPU usage of services
- Total disk I/O read rate of services
- Total disk I/O write rate of services

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Total used memory of services

Checking Static Resources (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Static Service Pool Configurations**.

Step 3 In the configuration group list, click a configuration group, for example, **default**.

Step 4 Check the system resource adjustment base values.

- **System Resource Adjustment Base** indicates the maximum volume of resources that can be used by each node in the cluster. If a node has only one service, the service exclusively occupies the available resources on the node. If a node has multiple services, all services share the available resources on the node.
- **CPU** indicates the maximum number of CPUs that can be used by services on a node.
- **Memory** indicates the maximum memory that can be used by services on a node.

Step 5 In **Chart**, view the metric data of the cluster service resource usage.

- You can click **Add Service to Chart** to include static service resource data for specific services. A maximum of 12 services can be added to the chart.
- For details about how to manage a chart, see **Viewing MRS Cluster Resource Monitoring Metrics**.

----End

Checking Static Resources (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**. In the **Resource** area, click **Configure Static Service Pool**.

Step 3 Click **Status**.

Step 4 Check the system resource adjustment base values.

- **System Resource Adjustment Base** indicates the maximum volume of resources that can be used by each node in the cluster. If a node has only one service, the service exclusively occupies the available resources on the node. If a node has multiple services, all services share the available resources on the node.
- **CPU(%)** indicates the maximum number of CPUs that can be used by services on a node.
- **Memory(%)** indicates the maximum memory that can be used by services on a node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Check the cluster service resource usage.

In the chart area, select **All services** from the service drop-down list box. The resource usage status of all services in the service pool is displayed.

Effective Configuration Group indicates the resource control configuration group currently in use by the cluster service. By default, the **default** configuration group is applied at all times, allowing the cluster service to use all CPUs and 70% of the node's memory.

Step 6 View the resource usage of a single service.

In the chart area, select a service from the service drop-down list box. The resource usage status of the service is displayed.

Step 7 You can set the interval for automatically refreshing the page.

Step 8 In the **Period** area, select a time range for viewing service resources.

Step 9 Click **View** to view the service resource data in the corresponding time range.

Step 10 Customize a service resource report.

1. Click **Customize** and select the service source indicators to be displayed.
2. Click **OK** to save the selected monitoring metrics for display.
3. Click **Clear** to cancel all selected monitoring metrics at once.

Step 11 Export a monitoring report.

Click **Export**. MRS Manager will generate a report about the selected service resources in a specified time of period. Save the report.

Click **View** to display curve charts of monitoring metrics for a specified period.

----End

8.13 Managing SQL Inspection Rules for an MRS Cluster

8.13.1 SQL Inspector Overview

New SQL engines are constantly emerging for big data. While contributing to increased diversity of big data solutions, they also lead to many issues, for example, the varying quality of SQL statements, increased difficulty in troubleshooting SQL-related performance issues, and excessive resource consumption by large queries.

Low-quality SQL statements may have unpredictable impacts on the performance and stability of data analytics platforms.

Notes and Constraints

- The SQL Inspector for Hive , Spark, HetuEngine, and ClickHouse is available only in MRS 3.3.0 or later.
- The SQL Inspector for Doris is available only in MRS 3.5.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Introduction to SQL Inspector

MRS offers the SQL Inspector that allows you to detect, warn about, intercept, or terminate large SQL queries and low-quality SQL statements for mainstream SQL engines, such as Hive, Spark, HetuEngine, ClickHouse, and Doris. The implementation of SQL Inspector does not require any service changes. Users do not need to change the way they submit SQL statements or the SQL syntax they are familiar with.

- You can configure SQL inspection rules on the UI that also allows you to query and modify the rules.
- During query execution and response, each SQL engine proactively inspects SQL statements based on the rules.
- You can choose one of the following actions for problematic SQL statements: Hint, Intercept, or Block. Hint means to log and warn the user about large SQL queries; Intercept means to prevent them from being executed; and Block means to terminate them while they are being executed. The system logs SQL inspection events in real time for auditing purposes. O&M engineers can analyze the logs, evaluate SQL statement quality on the live network, identify problematic statements, and take appropriate measures.

SQL inspection rules are classified into the following types:

- Static interception: displays hints or intercepts SQL statements based on SQL syntax.
- Dynamic interception: displays hints or intercepts SQL statements based on the rules regarding data table statistics and metadata information.
- Runtime blocking: terminates SQL statements that are being executed based on real-time system status information, such as CPU, memory, and I/O.

SQL statements that trigger static or dynamic interception rules are intercepted. Alternatively, the system gives hints for you to handle the statements properly. SQL statements that trigger runtime blocking rules are terminated.

SQL Inspection Rules and Constraints

- A SQL inspection rule can be associated with multiple SQL engines, and have different threshold parameters configured for each service.
- A SQL inspection rule can be associated with multiple tenants, and takes effect only for associated tenants.

8.13.2 Adding an MRS SQL Inspection Rule

You can add rules for specified tenants and SQL engines on FusionInsight Manager. The system will display hints on, intercept, or block SQL requests matched by the rules.

WARNING

Exercise caution when adding or modifying SQL inspection rules for a cluster, enabling rules, or configuring thresholds. Misconfigured rules may result in interruptions to upper- layer services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Adding a Rule

Step 1 Log in to MRS Manager as a user who has been assigned MRS Manager administrator permissions.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Click **Cluster** and choose **SQL Inspector**. The **SQL Inspector** page is displayed.

You can click **View Supported Rules** to view all SQL inspection rules supported by the current cluster.

Step 3 Click **Add Rule**. After the password of the current user is verified, the **Add Rule** page is displayed.

Step 4 Set the required parameters and click **OK**.

Parameter	Example Value	Description
Name	SQL_defenses01	Name of a SQL inspection rule
ID	static_0001	Rule ID For details about the meanings of specific rule IDs, see **Table 8-71**.
Tenant	default	Click **Add** to select the name of the tenant to which the current rule will be associated. If you need to add a new tenant, plan and create a cluster tenant by referring to **Managing MRS Cluster Tenants**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Services and Actions	-	Click **Add** to specify the SQL engine to which this rule will be associated with and set the threshold parameters of the rule. Each rule can be associated with one SQL engine. If you want to configure a rule for other SQL engines, add new rules. • **Service**: Select the SQL engine associated with the current rule. • If an SQL request meets the rule, the system performs the following operations: – **Hint**: Record logs and display a hint for handling the SQL request. If the rule has parameters, you need to configure the threshold. – **Intercept**: Intercept the SQL request that meets the rule. If the rule has parameters, you need to configure the threshold. – **Block**: Block the SQL request that meets the rule. If the rule has parameters, you need to configure the threshold. For static and dynamic SQL interception rules, **Hint** and **Block** operations are supported. Only the **Block** option is available for blocking rules.

Step 5 View the added prevention rule on the **SQL Defense** page. The rule takes effect dynamically.

To adjust the current rule, click **Modify** in the **Operation** column of the row that contains the target rule. After the user password is verified, you can modify rule parameters.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS SQL Inspection Rules

Table 8-71 MRS SQL inspection rules

ID	Description	Engine	Threshold	Example SQL Statement
static_0001	Check whether the number of occurrences of count(distinct) in the SQL statement exceeds the limit.	• Hive • Spark • HetuEngine	Number of occurrences of count(distinct) Recommended value: **10**	SELECT COUNT(DISTINCT deviceId), COUNT(DISTINCT collDeviceId) FROM table GROUP BY deviceName, collDeviceName, collCurrentVersion;
static_0002	Check whether the not in <subquery> statement is used in the SQL statement.	• Hive • Spark • HetuEngine	N/A	SELECT * FROM Orders o WHERE Orders.Order_ID not in (Select Order_ID FROM HeldOrders h where h.order_id = o.order_id);
static_0003	Check whether the number of joins in the SQL statement exceeds the limit.	• Hive • Spark • HetuEngine	Number of joins Recommended value: **20**	N/A

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ID	Description	Engine	Threshold	Example SQL Statement
static_0004	Check whether the number of the **union all** operators in a SQL statement exceeds the threshold.	• Hive • Spark • HetuEngine	Number of the **union all** operators in a statement Recommended value: **20**	select * from tables t1 union all select * from tables t2 union all select * from tables t3 union all select * from tables t4 union all select * from tables t5 union all select * from tables t6 union all select * from tables t7 union all select * from tables t8 union all select * from tables t9;
static_0005	The number of subquery nesting layers exceeds the limit.	• Hive • Spark • HetuEngine	Maximum number of nested subqueries Recommended value: **20**	select * from (with temp1 as (select * from tables) select * from temp1);
static_0006	Check whether the length of the SQL statement string exceeds the upper limit.	• Hive • Spark • HetuEngine	Length of the SQL string, in KB Recommended value: **10**	N/A
static_0007	Check whether the Cartesian product exists when multiple tables are associated.	• Hive • Spark • HetuEngine	N/A	select * from A,B;
static_0008	Check whether alter table update operation is performed at the cluster level (on cluster).	ClickHouse	N/A	alter table testtb1 on cluster default_cluster update price=10.0 where id='100'

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ID	Description	Engine	Threshold	Example SQL Statement
static_0009	Check whether alter table delete operation is performed at the cluster level (on cluster).	ClickHouse	N/A	alter table testtb1 on cluster default_cluster delete where id ='10'
static_0010	Check whether the alter table add column operation is performed at the cluster level (on cluster).	ClickHouse	N/A	alter table testtb1 on cluster default_cluster add column testc String
static_0011	Check whether the alter table drop column operation is performed at the cluster level (on cluster).	ClickHouse	N/A	alter table testtb1 on cluster default_cluster drop column testc
static_0012	Check whether the optimize final operation is performed at the cluster level (on cluster).	ClickHouse	N/A	optimize table testtb1 on cluster default_cluster final
static_0013	Check whether the drop table operation is performed at the cluster level (on cluster).	ClickHouse	N/A	drop table testtb1 on cluster default_cluster;
static_0014	Check whether the truncate table operation is performed at the cluster level (on cluster).	ClickHouse	N/A	truncate table testtb1 on cluster default_cluster;
dynamic_0001	Check whether the number of scanned files exceeds the limit.	• Hive • Spark • HetuEngine	Number of files that will be scanned or have been scanned Recommended value: **100,000**	SELECT ss_ticket_number FROM store_sales WHERE ss_ticket_number= 72291252 LIMIT 10;

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ID	Description	Engine	Threshold	Example SQL Statement
dynamic_0002	Check whether the number of partitions involved in operations (select, delete, update, and alter) on a table exceeds the upper limit.	• Hive • Spark • HetuEngine • ClickHouse	Number of partitions involved in the delete or alter operation Recommended value: **10,000**	DELETE FROM table_name WHERE column_name = value
dynamic_0003	When the right table of a join is a distributed table, check whether the data volume of the right table exceeds the upper limit.	ClickHouse	Number of rows in the right table when the join operation is performed. Recommended value: **100,000,000**	SELECT name, text FROM table_1 JOIN table_2 ON table_1.Id = table_2.Id
running_0001	Check whether the number of result rows returned by the Select statement to the client exceeds the upper limit.	• Hive • Spark • HetuEngine • ClickHouse	Number of rows in the query result Recommended value: **100,000**	select * from table
running_0002	Check whether the peak memory usage of the SQL statement exceeds the absolute value limit.	• Hive • Spark • HetuEngine • ClickHouse	Memory occupied by SQL running, in MB	N/A
running_0003	Check whether the running duration of the SQL statement exceeds the upper limit.	• Hive • Spark • HetuEngine • ClickHouse	SQL running duration threshold, in seconds	N/A

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

ID	Description	Engine	Threshold	Example SQL Statement
running_0004	The amount of data scanned by the SQL statements.	• Hive • Spark • HetuEngine • ClickHouse	Amount of data scanned by the SQL statement, in GB Recommended value: **10,240**	N/A

8.13.3 Configuring Hive SQL Inspection Rules

You can configure rules for Hive SQL inspection on FusionInsight Manager and configure rule parameters as you need.

Prerequisites

- The cluster client that contains the Hive service has been installed in the **/opt/hadoopclient** directory.
- The Hive service of the cluster is running properly.
- For a cluster with Kerberos authentication enabled, a user with Hive operation permissions has been created.

Notes and Constraints

- By default, SQL inspection rules need 5 seconds to take effect dynamically. After the queue is modified, it takes 10 minutes for Hive inspection rules to be reloaded.
- Interception and blocking rules will interrupt SQL tasks, so you need to set parameters of these rules properly based on the site requirements.
- For the rule dynamic_0001 (the number of files scanned by SQL statements exceeds the threshold), when the Spark and Tez engines reach the threshold, interception logs are printed in Yarn task logs and cannot be output on the Beeline client.
- Blocking rules have execution latency. For example, if the running_0004 rule is used and the threshold of the scanned data volume is 10 GB, the statement may be blocked when the data volume is 15 GB or higher due to the determination period and task concurrency.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Cluster** and choose **SQL Inspector**. The **SQL Inspector** page is displayed.

Step 3 Add an SQL inspection rule for Hive by referring to **Adding an MRS SQL Inspection Rule**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about the SQL inspection rules supported by the Hive SQL engine, see **MRS SQL Inspection Rules**.

For example, add a rule whose ID is **static_0001** to check whether count distinct appears more than two times in the SQL statement. If so, the system displays a hint.

Step 4 Log in to the node where the Hive client is installed and run the following command to switch to the client installation directory.

cd /opt/hadoopclient

Run the following command to set environment variables:

source bigdata_env

Run the following command to authenticate the current user. Skip this step if Kerberos authentication is disabled for the cluster (the cluster is in normal mode).

kinit *Component service user who has the Hive operation permission*

Step 5 Run the following command to log in to the Hive client:

beeline

Step 6 Run the following commands to create a table and import data to the table.

drop table if exists hivetb;

create table hivetb(a int,b int);

insert into hivetb select 1,11;

insert into hivetb select 2,22;

Step 7 Run the following SQL statement to check whether the current rule takes effect:

select count(distinct a),count(distinct b) from hivetb;

If the number of **count distinct** expressions in the statement exceeds the threshold configured in **Step 3**, the system displays the following message:

```
...
WARN  : STATIC_0001 The count(distinct X) times exceeds the limit : 2, current count distinct times : 2
...
```

If the operation set in the rule is **Block**, the statement fails to be executed and the following information is displayed:

```
...
Error: Error while compiling statement: FAILED: RuleException STATIC_0001 The count(distinct X) times exceeds the limit : 2, current count distinct times : 2 (state=42000,code=40000)
...
```

NOTE

- For more Hive SQL inspection rules, see **MRS SQL Inspection Rules**.
- You can also obtain the SQL inspection rules via logs which are stored in **/var/log/Bigdata/audit/hive/hiveserver/queryinfo.log**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.13.4 Configuring ClickHouse SQL Inspection Rules

You can configure rules for ClickHouse SQL inspection on MRS Manager and configure rule parameters as you need.

Prerequisites

- The cluster client that contains the ClickHouse service has been installed in the **/opt/hadoopclient** directory.
- The ClickHouse logical cluster is running properly.
- For clusters with Kerberos authentication enabled, you need to create a service user who has the permission to operate the ClickHouse table. For example, create a human-machine user **clickhouseuser**.
- A tenant for the ClickHouse service has been created and associated with the corresponding ClickHouse service user. For details, see **Creating an MRS Tenant**.

Notes and Constraints

- The default dynamic validity period of a rule is 1 minute.
- Interception and blocking rules will interrupt SQL queries, so you need to set parameters of these rules properly based on the site requirements.
- After configuring ClickHouse rules, you need to log in to the client again for the rules to take effect.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Cluster** and choose **SQL Inspector**. The **SQL Inspector** page is displayed.

Step 3 Add SQL inspection rules for ClickHouse by referring to **Adding an MRS SQL Inspection Rule**.

For details about the SQL inspection rules supported by the ClickHouse SQL engine, see **MRS SQL Inspection Rules**.

For example, add a rule with ID **static_0008** that triggers a prompt when an SQL statement performs a ClickHouse cluster- level table update.

Step 4 Log in to the node where the ClickHouse client is installed and run the following command to switch to the client installation directory.

cd /opt/hadoopclient

Run the following command to set environment variables:

source bigdata_env

Step 5 If the current cluster is in security mode (Kerberos authentication is enabled), run the following command to authenticate the current user. The current user must have the permission to create ClickHouse tables. If the current cluster is in normal mode (Kerberos authentication is disabled), skip this step.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

kinit *Component service user*

Example: **kinit clickhouseuser**

Step 6 Use the ClickHouse client to connect to the ClickHouse server.

Security mode

clickhouse client --host *IP address of the ClickHouseServer instance* **--port 9440 --secure**

Normal clusters:

clickhouse client --host *IP address of the ClickHouseServer instance*--**user** *Username* **--password --port 9000**

Enter the password.

Step 7 Run the following statements to create a data table:

CREATE DATABASE *cktest* **ON CLUSTER** *default_cluster*;

CREATE TABLE *cktest.test2* **ON CLUSTER** *default_cluster* **(`EventDate` DateTime, `CounterID` UInt32, `UserID` UInt32, `ver` UInt16) ENGINE = ReplicatedMergeTree('/clickhouse/tables/{shard}/***cktest/test2***', '{replica}') PARTITION BY toYYYYMM(EventDate) ORDER BY (EventDate, intHash32(UserID));**

CREATE TABLE *cktest.test2_dir* **ON CLUSTER** *default_cluster* **as** *cktest.test2* **ENGINE = Distributed(***default_cluster, cktest, test2***, rand());**

Step 8 Run the following command to insert data to the table:

insert into *cktest.test2* **values('2023-08-01',111,111,111);**

insert into *cktest.test2* **values('2023-08-02',222,111,111);**

Step 9 Run the following SQL statement for the created table to check whether the rule takes effect:

alter table *cktest.test2* **on cluster** *default_cluster* **update CounterID = toUInt32(222) where EventDate='2023-08-01' ;**

```
...
<Warning> SQLDefender: Distributed DDL ALTER UPDATE queries are undesirable.
...
```

If the operation set in the rule is **Intercept**, the statement fails to be executed and the following information is displayed:

```
...
DB::Exception: Distributed DDL ALTER TABLE UPDATE queries are undesirable..(QUERY_IS_PROHIBITED)
...
```

NOTE

For more ClickHouse SQL inspection rules, see **MRS SQL Inspection Rules**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

8.13.5 Configuring HetuEngine SQL Inspection Rules

You can configure rules for HetuEngine SQL inspection on MRS Manager and configure rule parameters as you need.

Prerequisites

- The cluster client that contains the HetuEngine service has been installed in the **/opt/hadoopclient** directory.
- The HetuEngine service and compute instances are running properly.
- If Kerberos authentication has been enabled for the cluster, you need to create a HetuEngine user and grant related permissions to the user. In addition, you need to use Ranger to assign the user the permission to manage databases, tables, and columns of the data source.

Notes and Constraints

- The default dynamic validity period of a rule is 5 minutes.
- Interception and blocking rules will interrupt SQL queries, so you need to set parameters of these rules properly based on the site requirements.
- Blocking rules are controlled by session-level parameters of the system. To configure blocking rules, service users must have the set session permission.
- For static rule **static_0003**, the total number of joins in queries does not include Semi joins and Anti joins.
- When prompt rules are configured for dynamic_0001 and dynamic_0002, prompt messages are recorded only in logs and are not displayed on the client.
- The client and server send asynchronous requests. For blocking rule **Running_0001**, after the server blocks the requests, the message "Query is gone " may be displayed on the client. In this case, you can view logs to check whether the requests are blocked.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Cluster** and choose **SQL Inspector**. The **SQL Inspector** page is displayed.

Step 3 Add SQL inspection rules for HetuEngine by referring to **Adding an MRS SQL Inspection Rule**.

For details about the SQL inspection rules supported by the HetuEngine SQL engine, see **MRS SQL Inspection Rules**.

For example, add a rule whose ID is **static_0001** to check whether count distinct appears more than two times in the SQL statement. If so, the system displays a hint.

Step 4 Log in to the node where the HetuEngine client is installed and run the following command to switch to the client installation directory:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cd /opt/hadoopclient

Run the following command to set environment variables:

source bigdata_env

Step 5 Log in to the HetuEngine client based on the cluster authentication mode.

- In security mode, run the following command to authenticate the user and log in to the HetuEngine client:

  kinit hetu_test

  hetu-cli --catalog hive --tenant default --schema default

- In normal mode, run the following command to log in to the HetuEngine client:

  hetu-cli --catalog hive --tenant default --schema default --user hetu_test

  hetu_test is a service user who has at least the tenant role specified by **--tenant** and cannot be an OS user.

Step 6 Check whether the current rule takes effect.

Run the following statement to create a table:

CREATE TABLE table1(id int, name varchar,rank int);

INSERT INTO table1 VALUES(10,'sachin',1),(45,'rohit',2),(46,'rohit',3),(18,'virat',4),(25,'dhawan',5);

Run the following statement to query data:

select count(distinct id),count(distinct id),count(distinct id),count(distinct id),count(distinct id),count(distinct id) from table1;

If the number of **count distinct** expressions in the statement exceeds the threshold configured in **Step 3**, the system displays the following message:

```
WARNING: Occurrence number of 'COUNT(DISTINCT XX)' (6) reaches the hint limitation (2)
```

- If the action set in the rule is **Intercept** or **Block**, the system may display the following message:
  ```
  Intercepted. Reason: Occurrence number of 'COUNT(DISTINCT XX)' (6) reaches the interception limitation (2)
  ```
- You can query HetuEngine SQL inspection details in logs stored in **hdfs://hacluster/hetuserverhistory/*Tenant*/coordinator/*application_ID*/*container_ID/yyyyMMdd*/server.log**.
- If JDBC secondary development requires warning information, add the following configuration to the JDBC application:
  ```
  statement = connection.prepareStatement(sql.trim());
  resultSet = statement.executeQuery();
  SQLWarning sqlWarning = statement.getWarnings();
  ```

----End

8.13.6 Configuring Spark SQL Inspection Rules

You can configure rules for Spark SQL inspection on FusionInsight Manager and configure rule parameters as you need.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- The cluster client that contains the Spark service has been installed in the **/opt/hadoopclient** directory.
- Spark is running properly.
- A tenant, for example, **sparkstatic1** has been added to **Tenant Resources**. For details, see **Creating an MRS Tenant**.
- For a cluster with Kerberos authentication enabled, a service user has been created, for example, user **sparkuser**. The user belongs to groups **hive**, **hadoop**, and **supergroup**, the primary group is **hive**, and the bound role is set to **sparkstatic1**.

Notes and Constraints

- The default dynamic validity period of a rule is 6 minutes.
- Only SQL jobs are supported.
- Spark SQL does not support SQL inspection in local mode.
- The SQL inspector cannot scan custom jobs submitted through spark-submit.
- Interception and blocking rules will interrupt SQL queries, so you need to set parameters of these rules properly based on the site requirements.
- The static rule static_0007 is not required for Spark because it has a Cartesian product restriction (controlled by **spark.sql.crossJoin.enabled**, which is **true** by default). If the spark parameter is **false**, the static_0007 rule will not take effect.
- Dynamic rules do work on carbon tables.
- The dynamic rule dynamic_0002 supports SELECT, ALTER TABLE ADD PARTITION and ALTER TABLE DROP PARTITION. If you use a batch deletion statement that contains judgment conditions, for example, ALTER TABLE DROP PARTITION (pt < 10), the statement will be intercepted before the dynamic_0002 rule because the number of partitions is limited by **spark.sql.dropPartitionsInBatch.limit**, which is defaulted to **1000**.
- Blocking rules have execution latency. For example, if the running_0004 rule is used and the threshold of the scanned data volume is 10 GB, the statement may be blocked when the data volume is 15 GB or higher due to the determination period and task concurrency.
- If a job does not exceed the threshold until the last several tasks are complete before the block action is triggered, the job cannot be canceled.
- Blocking rule running_0004: The SQL execution duration includes the execution duration on the Driver side and the job running duration. When the SQL execution is blocked on the Driver side, the job cannot be canceled even if the execution duration exceeds the value of interruption threshold. This problem may occur when INSERT OVERWRITE is performed on a large number of partitions where storage and compute are decoupled.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Click **Cluster** and choose **SQL Inspector**. The **SQL Inspector** page is displayed.

Step 3 Add an SQL inspection rule for Spark by referring to **Adding an MRS SQL Inspection Rule**.

For details about the SQL inspection rules supported by the Spark SQL engine, see **MRS SQL Inspection Rules**.

For example, add a rule whose ID is **static_0001** to check whether count distinct appears more than two times in the SQL statement. If so, the system displays a hint.

Step 4 Log in to the node where the Spark client is installed and run the following command to switch to the client installation directory.

cd /opt/hadoopclient

Run the following command to set environment variables:

source bigdata_env

source Spark/component_env

Step 5 Perform user authentication for clusters in security mode (Kerberos authentication enabled). Skip this step for clusters in normal mode (Kerberos authentication disabled).

kinit *Spark operation user*

Example:

kinit sparkuser

Enter the password as prompted and change the password upon your first login.

Step 6 Run the following command to log in to the spark-sql client:

cd opt/client/Spark/spark/bin

./spark-sql

Step 7 Run the following SQL statement on the client to check whether the current SQL inspection rule takes effect:

Run the following statement to create a table:

create table table1(id int, name string) stored as parquet

Run the following statement to query data:

select count(distinct id),count(distinct id),count(distinct id),count(distinct id),count(distinct id),count(distinct id) from table1;

If the number of **count distinct** expressions in the statement exceeds the threshold configured in **Step 3**, the system displays the following message:

```
WARNING:  static_0001 Occurrence num of 'COUNT(DISTINCT)'(6) reaches the hint threshold(2)
```

If the action set in the rule is **Intercept**, the following information is displayed:

```
Error in query: static_0001 Occurrence num of 'COUNT(DISTINCT)'(6) reaches the intercept threshold(2)
```

In Spark Beeline, you can obtain SQL inspection details from logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Yarn**. On the **Dashboard** page, click the link next to **ResourceManager WebUI** to enter the Yarn web UI.
2. Click the ID of the target application on the **All Applications** page. The application details page is displayed.
3. Click **Logs** of the application. On the displayed page, click **stdout** logs to view SQL inspection details.

NOTE

1. For more Spark SQL inspection rules, see **MRS SQL Inspection Rules**.
2. You can view query info in the **/opt/hadoopclient/Spark/spark/audit/query.log** path if you are using the Spark client. The log contains detailed running information and the corresponding SQL inspection information.

----End

8.13.7 Configuring Doris SQL Inspection

Scenarios

You can configure inspection rules for Doris SQL statements on MRS Manager and set rule parameters based on your needs.

Prerequisites

- The MySQL client is installed, and the node that will connect to the Doris database can communicate with the MRS cluster.
- The Doris FE and BE instances in the MRS cluster are functioning normally.

Notes and Constraints

- An SQL inspection rule is automatically applied in 5 minutes.
- Interception and blocking rules will interrupt SQL queries, so you need to set parameters of these rules properly based on the site requirements.
- This function is supported only in MRS 3.5.0 or later. In MRS 3.5.0, the static rule **static_0001** does not take effect for **count(distinct *xxx*)** operations in Doris.

  For example, the SQL statement is as follows:

  ```
  with t1 as (select count(distinct user_id) from example_tbl),t2 as (select count(distinct user_id) from example_tbl1),t3 as (select count(distinct error_code) from example_tbl2) select * from t1,t2,t3;
  select * from example_tbl where (select count(distinct user_id) from example_tbl) > 1 or (select count(distinct user_id) from example_tbl1) > 1 or (select count(distinct error_code) from example_tbl2) > 1;
  ```

Procedure

Step 1 Create a user with the Doris administrator permissions to connect to the Doris service.

1. Log in to FusionInsight Manager as user **admin**, choose **System** > **Permission** > **Role**, click **Create Role**, set the following parameters, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - **Role Name**: Enter a role name, for example, **dorisrole**.
  - Configure resource permission.

    In the **Configure Resource Permission** area, choose *Name of the desired cluster* > **Doris** and select **Doris Admin Privilege**.

2. Click **User**. If Kerberos authentication is enabled for the cluster (the cluster is in security mode), add a human-machine user. If Kerberos authentication is disabled for the cluster (the cluster is in normal mode), add a machine-machine user. Bind the user to the created role.
3. Log in to Manager as the newly-created user and change the initial password.

Step 2 Log in to FusionInsight Manager as a user with the **Manager_administrator** and **Manager_viewer** permissions and choose **Cluster** > **SQL Inspector**.

Step 3 Add an SQL inspection rule for Doris by referring to **Adding an MRS SQL Inspection Rule**.

For details about the rules supported by the Doris SQL engine, see **MRS SQL Inspection Rules**.

For example, add the **static_0001** rule while setting the threshold to **1** for the **Hint** action and **6** for the **Intercept** action. This rule detects a SQL statement that contains more than one **count distinct**.

Step 4 Log in to the node where MySQL is installed and connect to the Doris database.

If Kerberos authentication (security mode) has been enabled for the cluster, run the following commands to connect to the Doris database:

```
export LIBMYSQL_ENABLE_CLEARTEXT_PLUGIN=1
```

- Directly connect to the FE node to access the Doris database.

  mysql -u*Database login username* **-p -P***Connection port for FE queries* **-h***IP address of the Doris FE instance*

  Enter the password for logging in to the database.
- Connect the DBalancer service. DBalancer connects to the FE node to access the Doris database based on the configured policy.

  mysql -u*Database login user* **-p -P***TCP access port of DBalancer* **-h***IP address of the Doris DBalancer instance*

  Enter the password for logging in to the database.

Parameters are as follows:

- The database login user is the user created in **Step 1** and has the Doris administrator permissions.
- To obtain the query connection port of the Doris FE instance, log in to MRS Manager, choose **Cluster** > **Services** > **Doris** > **Configurations**, and check the **query_port** value of the Doris service.
- To obtain the TCP access port of the Doris FE instance, log in to MRS Manager, choose **Cluster** > **Services** > **Doris** > **Configurations**, and check the **balancer_tcp_port** value of the Doris service.
- To obtain the IP address of the Doris FE or DBalancer instance, log in to MRS Manager of the cluster, choose **Cluster** > **Services** > **Doris** > **Instances**, and check the service IP address of any FE or DBalancer instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You can also connect to the database using the MySQL connection software or Doris web UI.

Step 5 Check the configured SQL inspection rules.

```
use __internal_schema;
select * from sqldefend_rule;
```

The query result is as follows:

```
+-------------+----------+--------------------+--------+-----------------+------+------+------+
| rule_name   | is_group | group_or_user_name | action | threshold_value | p_1  | p_2  | p_3  |
+-------------+----------+--------------------+--------+-----------------+------+------+------+
| static_0001 |        1 | A                  |      1 |               3 | NULL | NULL | NULL |
| static_0001 |        1 | A                  |      2 |               2 | NULL | NULL | NULL |
+-------------+----------+--------------------+--------+-----------------+------+------+------+
```

Step 6 Create a database and then switch to that database.

```
create database test;
use test;
```

Step 7 Create tables **t1**, **t2**, and **t3**.

```
create table if not exists t1(id int) engine=olap distributed by hash(id);
create table if not exists t2(id int) engine=olap distributed by hash(id);
create table if not exists t3(id int) engine=olap distributed by hash(id);
```

Step 8 Check whether the **static_0001** rule has taken effect.

```
select count(distinct id) from t1 except select count(distinct id) from t2 intersect select count(distinct id) from t3 union all select count(distinct id) from t1 except select count(distinct id) from t2 intersect select count(distinct id) from t3;
```

The number of **count distinct** expressions in this statement exceeds the threshold configured in **Step 3**.

- The system displays the following hint:

  ```
  ... join count = 0, countDistinct = 6 ,not in query = false ,union count = 1, nested subquery = 0 ,cartesian join = false
  ```

- The interception information is as follows:

  ```
  ERROR ... detailMessage = static_0001 number of count(distinct) in the query 6 more than the allowed limit 3
  ```

If the action specified in the SQL inspection rule is **Block**, the system may display the following message:

```
ERROR ... detailMessage = running_0001 Num of result rows num reaches the fuse threshold(1000), so cancel the output
```

You can query Doris SQL inspection details in **/var/log/Bigdata/doris/fe/fe.log** and **/var/log/Bigdata/audit/doris/fe/fe.audit.log**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9 MRS Cluster O&M

9.1 Cluster O&M

Account Maintenance Suggestions

It is recommended that the administrator conduct routine checks on the accounts. The check covers the following items:

- Check whether the accounts of the OS, Manager, and each component are necessary and whether temporary accounts have been deleted.
- Check whether the permissions of the accounts are appropriate. Different administrators have different rights.
- Check and audit the logins and operation records of all types of accounts.

Password Maintenance Suggestions

Accessing portal requires identity authentication. The complexity and validity period of an account password must meet your security requirements.

Refer to the following suggestions to maintain passwords:

- Assign dedicated personnel to keep OS passwords.
- Use passwords that meet certain strength requirements, such as minimum password length or mixing of letter cases.
- Encrypt passwords before transferring them, and do not transfer them via email.
- Encrypt passwords for storage.
- Remind enterprise users to change passwords during system handover.
- Change passwords periodically.

Log Maintenance Suggestions

Operation logs help discover exceptions such as illegal operations and login by unauthorized users. The system records important operations in logs. You can use operation logs to locate problems.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Checking Logs Regularly

  Check system logs periodically and handle exceptions such as unauthorized operations or logins in a timely manner.

- Backing Up Logs Regularly

  Audit logs provided by Manager and clusters record user activity and operation information. You can export audit logs from Manager. If there are too many audit logs in the system, you can configure dump parameters to dump audit logs to a specified server to ensure that the cluster nodes disk space is sufficient.

- Maintenance Owner

  Network monitoring engineers and system maintenance engineers

Manager Routine Maintenance

To ensure long-term and stable running of the system, administrators or maintenance engineers need to periodically check items listed in the following table and rectify the detected faults based on the check results. It is recommended that administrators or engineers record the result in each task scenario and sign off based on the enterprise management regulations.

Table 9-1 Routine maintenance check items

Routine Maintenance Frequency	Role	Check Item
Daily	Check the cluster service status.	• Check whether the running status and configuration status of each service are normal and whether the status icons are green. • Check whether the running status and configuration status of the role instances in each service are normal and whether the status icons are green. • Check whether the active/standby status of role instances in each service can be properly displayed. • Check whether the dashboard of the services and role instances can be displayed properly.
	Check the cluster host status.	• Check whether the running status of each host is normal and whether the status icon is green. • Check the current disk usage, memory usage, and CPU usage of each host. Check whether the current memory usage and CPU usage are increasing.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Routine Maintenance Frequency	Role	Check Item
	Check the cluster alarm information.	Check whether alarms were generated for unhandled exceptions on the previous day, including alarms that were automatically cleared.
	Check the cluster audit information.	Check whether critical and major operations are performed on the previous day and whether the operations are valid.
	Check the cluster backup status.	Check whether OMS, LDAP, DBService, and NameNode have been automatically backed up on the previous day.
	View the health check result.	Perform a health check on Manager and download the health check report to check whether the current cluster is abnormal. You are advised to enable the automatic health check, export the latest cluster health check result, and repair unhealthy items based on the result.
	Check the network communication.	Check the cluster network status and check whether the network communication between nodes is delayed.
	Check the storage status.	Check whether the total data storage volume of the cluster increases abruptly. • Check whether the disk usage is close to the threshold. If yes, locate the causes. For example, check whether the junk data or cold data left by services needs to be cleared. • Check whether disk partitions need to be expanded based on the service growth trend.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Routine Maintenance Frequency	Role	Check Item
	Check logs.	• Check whether there are failed or unresponsive MapReduce and Spark tasks. Check the **/tmp/logs/${username}/logs/${application id}** log file in HDFS and rectify faults. • Check Yarn task logs, view the logs of failed and unresponsive tasks, and delete duplicate data. • Check the worker logs of Storm. • Back up logs to the storage server.
Weekly	User management	Check whether the user password is about to expire and notify the user of changing the password. To change the password of a machine-machine user, you need to download the keytab file again.
	Analyze alarms.	Export and analyze alarms generated in a specified period.
	Scan disks.	Check the disk health status. You are advised to use a dedicated disk check tool.
	Collect statistics on storage.	Check in batches whether the disk data of cluster nodes is evenly stored, filter out the disks whose data increases significantly or is insufficient, and check whether the disks are normal.
	Record changes.	Arrange and record the operations on cluster configuration parameters and files to provide reference for fault analysis and handling.
Monthly	Analyze logs.	• Collect and analyze hardware logs of cluster node servers, such as BMC system logs. • Collect and analyze the OS logs of the cluster node servers. • Collect and analyze cluster logs.
	Diagnose the network.	Analyze the network health status of the cluster.
	Manage hardware.	Check the equipment room environment and clean the devices.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.2 Logging In to an MRS Cluster

9.2.1 Checking MRS Active/Standby Management Nodes

Management nodes in an MRS cluster are the nodes where OMS is located. The node where the active OMS is located is the active management node, and the node where the standby OMS is located is the standby management node.

OMS nodes are usually deployed in active/standby mode, which are the **master1** and **master2** nodes in the Master node group of the cluster. When the active OMS node is faulty or stopped, the system triggers an active/standby switchover where the standby OMS node becomes active. Namely, not all **master1** nodes are active management nodes.

Some O&M scripts and commands of the MRS cluster need to be executed or can be executed only on the active management node. To determine the active and standby management nodes, log in to any management node or FusionInsight Manager.

Notes and Constraints

The FusionInsight Manager only displays the active/standby status of management nodes of MRS 3.*x* or later clusters.

Querying the Active and Standby OMS Nodes on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts**.

Step 3 In the host list, view and record the IP addresses of the active and standby management nodes.

- If a host name starts with ★, it is the active management node (active OMS node). View and record the value of **Management IP Address** in the row containing the active node.
- If a host name starts with ☆, it is the standby management node (standby OMS node). View and record the value of **Management IP Address** in the row containing the standby node.

----End

Querying the Active and Standby OMS Nodes on the MRS Console

Step 1 Obtain information about the Master nodes of the MRS cluster.

1. Log in to the MRS console and choose **Active Clusters**.
2. Click the name of the target cluster to access its details page.
3. Click **Synchronize** next to **IAM User Sync** to synchronize IAM users. If IAM users have been synchronized, skip this step.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. On the **Nodes** tab, view Master node names. The node that contains **master1** in its name is the Master1 node. The node that contains **master2** in its name is the Master2 node.

   If a node contains ★ in its name, it is the active management node (active OMS node). View and record the IP address of the node.

   If a node contains ☆ in its name, it is the standby management node (standby OMS node). View and record the IP address of the node.

   If there is no management node with the active/standby flag, go to **Step 2**.

Step 2 Check the active and standby management nodes in a cluster.

1. Log in to the Master1 node as the **root** user. For details, see **Logging In to an MRS Cluster Node**.
2. Switch to user **omm**.
   ```
   su - omm
   ```
3. Check the active/standby status of the management nodes.
   - For **MRS 2.*x*** and earlier versions, run the following command:
     ```
     sh ${BIGDATA_HOME}/om-0.0.1/sbin/status-oms.sh
     ```
   - For **MRS 3.*x*** and later versions, run the following command:
     ```
     sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh
     ```

   In the following command output, the node whose **HAActive** is **active** is the active management node (**mgtomsdat-sh-3-01-1** in the following example), and the node whose **HAActive** is **standby** is the standby management node (**mgtomsdat-sh-3-01-2** in the following example).

   ```
   Ha mode
   double
   NodeName            HostName                 HAVersion         StartTime
   HAActive            HAAllResOK          HARunPhase
   192-168-0-30        mgtomsdat-sh-3-01-1      V100R001C01       20xx-11-18 23:43:02
   active              normal              Actived
   192-168-0-24        mgtomsdat-sh-3-01-2      V100R001C01       20xx-11-21 07:14:02
   standby             normal              Deactived
   ...
   ```

   If the Master1 node to which you have logged in is the standby management node and you need to log in to the active management node, run the following command:

   ssh *IP address of Master2 node*

----End

9.2.2 Logging In to an MRS Cluster Node

This section describes how to log in to a node in an MRS cluster and manage or maintain the cluster. You can log in to a cluster node using VNC provided on the Elastic Cloud Server (ECS) console. Alternatively, you can log in to a node using SSH key or password. If the local environment cannot directly access the node VPC, you need to bind an elastic IP address (EIP) to the node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-2 MRS cluster node access methods

Access Method	Access Scenario	Restrictions
SSH	Routine node access, to install cluster clients, check role configuration files or log files of cluster components, and query data files required by upper-layer cluster jobs and so on.	You need to bind an EIP to the ECS and add a security group rule to allow inbound traffic. If the ECS is accessed through a private network, for example, through a VPN or Direct Connect connection, you do not need to bind an EIP to the ECS.
VNC (Remote Login from ECS Console)	To remotely log in to an ECS for maintenance in case of emergency.	No EIPs are required.

Prerequisites

- To log in to an MRS cluster node using SSH, ensure that an EIP is bound to the node in advance if your local environment cannot directly reach the MRS cluster network.
- Before logging in to an MRS cluster node using SSH, make sure the cluster's security group includes an inbound rule that allows the access.

  In the security group rule, the source address is the client IPv4 address/32 (or client IPv6 address/128), and the port number is 22.

Logging In to an MRS Cluster Node

You can log in to the node using SSH or remotely log in to the node using VNC on ECS console. The node login method depends on the node password type selected during MRS cluster creation.

Logging In to a Node Using SSH

Logging In to a Node on a Local Windows OS

To log in to a cluster node on a local Windows OS, perform the following operations:

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters** and click the name of the target cluster to access cluster details page.

Step 3 On the **Nodes** tab, click the name of the target node to log in to the ECS console.

To log in to the active or standby management node of the MRS cluster, see **Checking MRS Active/Standby Management Nodes**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Click the EIPs tab, click **Bind EIP** to bind an EIP to the ECS, record the IP address, and click **OK**. If an EIP has been bound, skip this step.

Step 5 Run the SSH login software, such as PuTTY.

Step 6 Click **Session**.

1. **Host Name (or IP address)**: Enter the EIP bound to the ECS.
2. **Port**: Enter **22**.
3. **Connection Type**: Select **SSH**.
4. **Saved Sessions**: Task name, which can be clicked for remote connection when you use PuTTY next time

Step 7 Click **Window** and select **UTF-8** for **Remote character set:** in **Translation**.

Step 8 Click **Open** to log in to the ECS.

If you log in to the ECS for the first time, PuTTY displays a security warning dialog box, asking you whether to accept the ECS security certificate. Click **Yes** to save the certificate to your local registry.

Step 9 After the SSH connection to the ECS is set up, enter the username and password as prompted to log in to the ECS.

The username is root, and the password is the one you set when creating the MRS cluster.

----End

Logging In to the Node on a Local Linux OS

To log in to a cluster node on a local Linux OS, bind an EIP to the ECS by referring to **Step 1** to **Step 4**, and then run the following command in the CLI to log in to the ECS:

```
ssh EIP bound to the ECS
```

Logging In to a Node Using an SSH Key Pair

Logging In to a Node on a Local Windows OS

To log in to the Linux ECS from local Windows, perform the operations described in this section. The following procedure uses PuTTY as an example to log in to the ECS.

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters** and click the name of the target cluster to access cluster details page.

Step 3 On the **Nodes** tab, click the name of the target node to log in to the ECS console.

To log in to the active or standby management node of the MRS cluster, see **Checking MRS Active/Standby Management Nodes**.

Step 4 Click the **EIPs** tab, click **Bind EIP** to bind an EIP to the ECS, and record the EIP. If an EIP has been bound to the ECS, skip this step.

Step 5 Check whether the private key file has been converted to **.ppk** format.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 10**.
- If no, go to **Step 6**.

Step 6 Run the SSH login software, such as PuTTY.

Step 7 In the **Actions** area, click **Load** and import the private key file you used during ECS creation.

Ensure that the private key file is in the format of **All files (*.*)**.

Step 8 Click **Save private key**.

Step 9 Save the converted private key,

for example, **kp-123.ppk**, to a local directory.

Step 10 Run PuTTY.

Step 11 Choose **Connection** > **Data**. Enter the image username in **Auto-login username**.

The default username for logging in to the MRS cluster node image is **root**.

Step 12 Choose **Connection** > **SSH** > **Auth**. In the last configuration item **Private key file for authentication**, click **Browse** and select the private key file in **.ppk** format or the private key file saved in **Step 9**.

Step 13 Click **Session**.

1. **Host Name (or IP address)**: Enter the EIP bound to the ECS.
2. **Port**: Enter **22**.
3. **Connection Type**: Select **SSH**.
4. **Saved Sessions**: Task name, which can be clicked for remote connection when you use PuTTY next time

Step 14 Click **Open** to log in to the ECS.

If you log in to the ECS for the first time, PuTTY displays a security warning dialog box, asking you whether to accept the ECS security certificate.

Click **Yes** to save the certificate to your local registry.

----End

Logging In to the ECS from Local Linux

To log in to the Linux ECS from local Linux, perform the operations described in this section.

The following procedure uses private key file **kp-123.pem** as an example to log in to the ECS. The name of your private key file may differ.

Step 1 On the Linux CLI, run the following command to change operation permissions:

```
chmod 400 /path/kp-123.pem
```

In the preceding command, **path** indicates the path for storing the private key file.

Step 2 Run the following command to log in to the ECS:

```
ssh -i /path/kp-123.pem Default username@EIP
```

For example, if the default username is **root** and the EIP is *xxx.xxx.xxx.xxx*, run the following command:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
ssh -i /path/kp-123.pem root@xxx.xxx.xxx.xxx
```

In the preceding command:

- *path* indicates the path where the key file is saved.
- *EIP* indicates the EIP bound to the ECS.
- The image username is **root** for cluster nodes.

----End

Logging In to a Node Using VNC

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters** and click the name of the target cluster to access cluster details page.

Step 3 On the **Nodes** tab, click the name of the target node to log in to the ECS console.

To log in to the active or standby management node of the MRS cluster, see **Checking MRS Active/Standby Management Nodes**.

Step 4 Click **Remote Login** in the upper right corner. In the displayed dialog box, click **Log In** in the **VNC Login** area.

Step 5 Enter the username and password for logging in to the Master node as prompted.

1. If you select **Password** for **Login Mode**, you need to enter **root** in **Username** and the password you set during cluster creation in **Password**.
2. If you select **Key Pair** for **Login Mode** when creating a cluster, perform the following operations to log in to the cluster:
   a. After the cluster is created, assign an EIP and bind it to the Master node of the cluster. For details, see **Virtual Private Cloud** > **User Guide** > **Elastic IP Address** > **Assigning an EIP and Binding It to an ECS**.
   b. Remotely log in to the Master node in SSH mode as user **root** using the key file.
   c. Run the following command to set the password for user **root**:
      ```
      passwd root
      ```
   d. Go back to the login interface, and enter **root** and the password set in **Step 5.2.c** to log in to the node.

----End

Helpful Links

To periodically change the password for the MRS cluster user to improve system security, see **Changing the Passwords for System Users of an MRS Cluster**.

9.3 Viewing MRS Cluster Monitoring Metrics

9.3.1 Viewing MRS Cluster Resource Monitoring Metrics

MRS cluster nodes are classified into management nodes, control nodes, and data nodes. The change trends of key host monitoring metrics on each type of node

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

can be calculated and displayed as curve charts in reports based on the customized periods. MRS cluster metrics are monitored periodically. The average historical monitoring interval is about 5 minutes.

You can view the overall resource usage of a cluster on the MRS console or MRS Manager. The monitoring metrics of clusters, components, and nodes are supported. For details about how to view the metrics of components and nodes, see **Viewing MRS Cluster Component Monitoring Metrics** and **Viewing MRS Node Resource Monitoring Metrics**.

Viewing Monitoring Information on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the **Dashboard** tab, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.

Step 4 After the synchronization is complete, you can view the cluster monitoring metric report on the right of the page.

Step 5 In time range area, specify a period to view monitoring data.

Step 6 Customize a monitoring report.

1. Click **Customize** and select monitoring metrics to display.
2. Click **OK** to save the selected monitoring metrics for display.
3. Click **Deselect All** to deselect all metrics.

Step 7 Export a monitoring report.

Monitoring metric reports can be exported only for versions earlier than MRS 3.*x*.

1. Select a period.
2. Click **Export**. MRS will generate a report about the selected monitoring metrics in a specified time range. Save the report.

----End

Viewing Monitoring Information on Manager (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Homepage**.

Step 3 In the upper right corner of the chart area, click ▼ and choose **Customize** from the displayed menu.

NOTE

Monitoring data of the past 1 hour is displayed at an interval of 5 minutes. After you enter the **Real-time Monitoring** page, you can view that real-time monitoring data is displayed on the right of the monitoring chart at an interval of 5 minutes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 In the navigation pane, select an entity and some monitoring metrics.

Step 5 Click **OK** to view the metric details.

Step 6 Export a monitoring report.

- Exporting all monitoring data
  a. On the **Homepage**, select a time range in the upper right corner of the chart area, for example, **1w**.
     Real-time data is displayed by default, which cannot be exported. You can click to customize a time range.
  b. In the upper right corner of the chart area, click and choose **Export** from the displayed menu.
- Exporting monitoring data of a specified monitoring item
  a. On the **Homepage**, click in the upper right corner of any monitoring report pane in the chart area of the target cluster.
  b. Select a time range to obtain monitoring data, for example, **1w**.
     Real-time data is displayed by default, which cannot be exported. You can click to customize a time range.
  c. Click **Export**.

The interval on the horizontal axis of the chart varies depending on the time period you specify. Data monitoring rules are as follows:

- If the disk usage of the partition where GaussDB is deployed exceeds 80%, real-time monitoring data and monitoring data whose interval is 5 minutes will be deleted.
- **Storage resources (HDFS) in Tenant Resources (0 to 300 hours)**: The interval is 1 hour. The cluster must have been installed for at least 1 hour, and monitoring data of a maximum of 3 months is saved.

For clusters of versions earlier than MRS 3.3.0:

- **0 to 25 hours**: The interval is 5 minutes. The cluster must have been installed for at least 10 minutes, and monitoring data of a maximum of 15 days is saved.
- **25 to 150 hours**: The interval is 30 minutes. The cluster must have been installed for at least 30 minutes, and monitoring data of a maximum of 3 months is saved.
- **150 to 300 hours**: The interval is 1 hour. The cluster must have been installed for at least 1 hour, and monitoring data of a maximum of 3 months is saved.
- **300 hours to 300 days**: The interval is 1 day. The cluster must have been installed for at least 1 day, and monitoring data of a maximum of 6 months is saved.
- **Over 300 days**: The interval is 7 days. The cluster must have been installed for more than 7 days, and monitoring data of a maximum of 1 year is saved.

For clusters of MRS 3.3.0 and later versions:

- **0 to 21 hours and 20 minutes**: The interval is 5 minutes. The cluster must have been installed for at least 10 minutes, and monitoring data of a maximum of 90 days is saved.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **21 hours and 20 minutes to 128 hours**: The interval is 30 minutes. The cluster must have been installed for at least 30 minutes, and monitoring data of a maximum of 90 days is saved.
- **128 to 256 hours**: The interval is 1 hour. The cluster must have been installed for at least 1 hour, and monitoring data of a maximum of 90 days is saved.
- **256 hours to 256 days**: The interval is 1 day. The cluster must have been installed for at least 1 day, and monitoring data of a maximum of 90 days is saved.

----End

Viewing Monitoring Information on Manager (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** on MRS Manager.

Step 3 In **Period**, you can specify a period to view monitoring data. The options are as follows:

Real-time, last 3 hours, last 6 hours, last 24 hours, last 1 week, last 1 month, last 3 months, last 6 months, and custom time range

Step 4 Click **View** to view monitoring data in a specified period.

- You can view **Health Status** and **Roles** of each service on the **Service Summary** page of MRS Manager.
- Click  above the curve chart to view details about a metric.

Step 5 Customize a monitoring report.

Click **Customize** and select monitoring metrics to be displayed on MRS Manager. Click **OK** to save the selected monitoring metrics for display. Click **Clear** to cancel all the selected monitoring metrics in a batch.

MRS Manager supports a maximum of 14 monitoring metrics, but at most 12 customized monitoring metrics can be displayed on the page.

Step 6 Set an automatic refresh interval or click  for an immediate refresh. If you select **Full Screen**, the **Dashboard** window will be maximized.

The following refresh interval options are supported:

- **Refresh every 60 seconds**
- **Refresh every 120 seconds**
- **Stop refreshing**

Step 7 Export a monitoring report.

Select a period. Click **Export**. MRS Manager will generate a report about the selected monitoring metrics in a specified time of period. Select a directory to save the report.

To view the curve charts of monitoring metrics in a specified period, click **View**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

On MRS Manager, you can query the top, bottom, and average value curves of key service and host monitoring metrics, that is, the resource distribution information. MRS Manager allows you to view the monitoring data of the last hour.

You can also modify the resource distribution on MRS Manager to display both the top and bottom value curves in service and host resource distribution figures.

Resource distribution of some monitoring metrics is not recorded.

- View the resource distribution of service monitoring metrics.
  a. On MRS Manager page, click **Services**.
  b. Select the target service from the service list.
  c. Click **Resource Distribution**.

     Select key metrics of the service from **Metric**. MRS Manager displays the resource distribution of the metrics in the last hour.
- View the resource distribution of host monitoring metrics.
  a. On MRS Manager, click **Hosts**.
  b. Click the name of the specified host in the host list.
  c. Click **Resource Distribution**.

     Select key metrics of the host from **Metrics**. MRS Manager displays the resource distribution of the metrics in the last hour.
- Configure resource distribution.
  a. On MRS Manager, click **System**.
  b. In **Configuration**, click **Configure Resource Contribution Ranking** under **Monitoring and Alarm**.
  c. Change the number of resources to be displayed. The sum of the maximum value and minimum value of resource distribution cannot be greater than 5.

     Set **Number of Top Resources** to the number of top values.

     Set **Number of Bottom Resources** to the number of bottom values.
  d. Click **OK** to save the configurations.

----End

9.3.2 Viewing MRS Cluster Component Monitoring Metrics

You can manage the following status and metrics of all components (including role instances) on the MRS console: Status information includes operation, health, configuration, and role instance status. Indicator information includes key monitoring indicators for each component.

Viewing the Monitoring Information on the MRS Management Console

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters** and click a cluster name to go to the cluster details page.

Step 3 On the **Dashboard** tab, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On the MRS cluster details page, click **Components**.

Step 5 View component monitoring information.

1. Click a specified service in the list to view its status and metric information.
2. Select and view component-level monitoring metrics.
   a. In the **Charts** area, click **Customize** to customize service monitoring metrics.
   b. In **Period** area, select a time of period and click **View** to view the monitoring data within the time period.

Step 6 View role instance monitoring information.

1. In the component list, click a service name.
2. Click **Instance** to view the status of each role instance in the component.

   You can filter all instances of the same role in the upper right corner of the list. You can set search criteria in the role search area by clicking **Advanced Search**, and click **Search** to view specified role information. You can click **Reset** to reset the search criteria. Fuzzy search is supported.
3. Click the target role instance to view its status and metric information.
4. Customize and view monitoring graphs.
   a. In the **Charts** area, click **Customize** to customize service monitoring metrics.
   b. In **Period** area, select a time of period and click **View** to view the monitoring data within the time period.

----End

Viewing Component Monitoring Information on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Go to the **Service Management** page.

- For MRS 3.x and later, choose **Cluster** > **Services**.
- For MRS 2.x and earlier, click **Services**.

Step 3 View component monitoring information.

1. Click a specified service in the list to view its status and metrics.
2. Customize and export monitoring charts.
   a. In the **Chart** tab, click ⌄ and **Customize** to customize service monitoring metrics.
   b. Set a time range and click ⌄ and **Export** to export the monitoring data.
   c. Click the help button next to a metric name to view the detailed help information about the metric.

Step 4 View role instance monitoring information.

1. Click the service name.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Click **Instances** to view the role status.
3. Click the target role instance to view its status and metric information.
4. Customize and export monitoring charts.
   a. In the **Chart** tab, click and **Customize** to customize metric chart.
   b. Select a time range. The monitoring data within the time range is displayed.
   c. Click the help button next to a metric name to view the detailed help information about the metric.

----End

Component Resource Monitoring Summary

NOTE

This function is supported only in MRS 3.x and later versions.

Log in to FusionInsight Manager and choose **Cluster** > **Services** > *Target service*. Click **Resource**. The resource monitoring page is displayed.

Some services in the cluster provide service-level resource monitoring metrics. By default, the monitoring data of the latest 12 hours is displayed. You can click to set a time range. You can click to export the corresponding report information. If a monitoring item has no data, the report cannot be exported. The following table lists the services and monitoring items that support resource monitoring.

Table 9-3 Service resource monitoring

Service	Monitoring Metric	Description
HDFS	Resource Usage (by Tenant)	• Collects statistics on HDFS resource usage by tenant. • Views the metrics **Capacity** or **Number of File Objects**.
	Resource Usage (by User)	• Collects statistics on HDFS resource usage by user. • Views the metrics **Used Capacity** or **Number of File Objects**.
	Resource Usage (by Directory)	• Collects statistics on HDFS resource usage by directory. • Views the metrics **Used Capacity** or **Number of File Objects**. • You can click to configure space monitoring. Alternatively, you can specify an HDFS file system directory for monitoring.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Monitoring Metric	Description
	Resource Usage (by Replica)	• Collects statistics on HDFS resource usages by replica count. • Views the metrics **Used Capacity** or **File Count**.
	Resource Usage (by File Size)	• Collects statistics on HDFS resource usages by file size. • Views the metrics **Used Capacity** or **File Count**.
	Recycle Bin (by User)	• Collects statistics on the usage of the HDFS recycle bin by user. • Views the metrics **Recycle Bin Capacity** or **Number of File Objects**.
	Operation Count	• Collects the number of operations in HDFS.
	Automatic Balancer	Collects statistics on the execution speed of HDFS automatic balancer and the total capacity of the current balancer migration.
	NameNode RPC Open Connections (by User)	Displays the number of connections of each user in the Client RPC requests connected to NameNodes.
	Slow DataNodes	Displays DataNode that transmits or processes data slowly in the cluster.
	Slow Disks	Displays the disk that processes data slowly on the DataNode in the cluster.
HBase	Operation Requests in Tables	Displays the number of PUT, DELETE, GET, SCAN, INCREMENT, and APPEND operation requests in all tables on all RegionServers.
	Operation Requests on RegionServers	Displays the number of PUT, DELETE, GET, SCAN, INCREMENT, and APPEND operation requests and number of all operation requests in RegionServer.
	Operation Requests for Service	Displays the number of PUT, DELETE, GET, SCAN, INCREMENT, and APPEND operation requests in all regions on RegionServers.
	HFiles on RegionServers	Displays the number of HFiles in all RegionServers.
Hive	HiveServer2-Background-Pool Threads (by IP)	Displays the number of HiveServer2-Background-Pool threads of top users. These threads are measured and displayed in a measurement period.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Monitoring Metric	Description
	HiveServer2-Handler-Pool Threads (by IP)	Displays the number of HiveServer2-Handler-Pools of top users collected and displayed in a period.
	Used MetaStore Number (by IP)	Collects statistics on and displays the MetaStore usage of top users in a period.
	Number of Hive jobs	Displays the number of user-related jobs collected by Hive in a period.
	Number of Files Accessed in the Split Phase	Displays the number of files accessed by the underlying file storage system (HDFS by default) in the Split phase in a period.
	Hive Basic Operation Time	Collects time for creating a directory (mkdirTime), creating a file (touchTime), writing a file (writeFileTime), renaming a file (renameTime), moving a file (moveTime), deleting a file (deleteFileTime), and deleting a directory (deleteCatalogTime) in a period of time.
	Table Partitions	Displays the number of partitions in all Hive tables, which is displayed in the following format: *database # table name*, *number of table partitions*.
	HQL Map Count	Collects statistics on HQL statements executed in a period and the number of Map statements invoked during the execution. The displayed information includes users, HQL statements, and the number of Map statements.
	HQL Access Statistics	Displays the number of HQL access times in a period.
Kafka	Kafka Disk Usage Distribution	Displays the disk usage distribution statistics of the Kafka cluster.
Spark/ Spark2x	HQL Access Statistics	Collects HQL access statistics in a period, including the username, HQL statement, and HQL statement execution times.
Yarn	Used resources (by Task)	• Displays the number of CPU cores and memory used by a task. • Views the metrics **By memory** or **By CPU**.
	Resource Usage (by Tenant)	• Displays the number of CPU cores and memory used by a tenant. • Views the metrics **By memory** or **By CPU**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Monitoring Metric	Description
	Resource usage ratio (by Tenant)	• Displays the ratio of the number of CPU cores to the memory used by a tenant. • Views the metrics **By memory** or **By CPU**.
	Task Duration Ranking	Displays Yarn tasks sorted by time consumption.
	ResourceManager RPC Open Connections (by User)	Displays the number of client RPC connections to ResourceManager by user.
	HBase Operation Count	Collects statistics on the number and proportion of operations corresponding to each Yarn operation type.
	Ranking of Tasks in a Queue by Resource Usage	• Displays the resources consumed by the tasks running in a queue after the queue (tenant) is selected on the GUI. • Views the metrics **By memory** or **By CPU**.
	Ranking of Users in a Queue by Resource Usage	• Displays the resources consumed by the users who are running tasks in the queue after a queue (tenant) is selected on the GUI. • Views the metrics **By memory** or **By CPU**.
ZooKeeper	Used Resources (By Second-Level Znode)	• Displays the ZooKeeper level-2 znode resource status. • Views the metrics **By Znode quantity** or **By capacity**.
	Number of Connections (by Client IP Address)	Displays the ZooKeeper client connection resource status.

9.3.3 Viewing MRS Node Resource Monitoring Metrics

You can view the status and metrics of each node in the MRS cluster in real time to see the current resource usage.

Viewing Host Monitoring Charts

Operations on the MRS management console:

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 On the **Dashboard** page, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.

Step 4 Click **Nodes** and expand the node group information to view the status of all hosts.

Step 5 The host list of a group contains the **Node Name/Resource ID**, **IP**, **Status**, **Specifications**, **Disks**, and **AZ**.

Step 6 Click the target node in the list to view its status and metric information.

Step 7 Click the **Monitoring** tab to view the monitoring charts of the current node.

----End

Operations on Manager:

- MRS 3.x and later:
  a. Log in to FusionInsight Manager of the MRS cluster.
     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
  b. Click **Hosts** to view the host list.
  c. In the host list, click the specified host name to view the host overview information.
     The host details page contains the basic information area, disk status area, instance list area, and monitoring charts.
  d. Click the **Chart** tab to view the full monitoring chart information about the host.
     The **Chart** page displays all monitoring charts of the host.
- MRS 2.x and earlier:
  a. Log in to FusionInsight Manager of the MRS cluster.
     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
  b. Click **Hosts** to view the status of all hosts.
  c. Click the target host in the host list to view its status and metric information.
  d. Customize and export monitoring charts.
     i. In the **Charts** area, click **Customize** to customize service monitoring metrics.
     ii. In **Period** area, select a time of period and click **View** to view the monitoring data within the time period.
     iii. Click **Export** to export the displayed metrics.

Viewing Host Resource Overview

This function is only available in MRS 3.*x* or later.

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **Hosts** > **Resource Overview** to view the resource distribution information.

In the **Distribution** tab, you can view the resource distribution of the cluster. By default, the monitoring data of the last hour is displayed. You can click to set a time range.

- Click **Select Metric** to specify the metrics you want to check. After you select a metric, the host distribution in each range of the metric is displayed.
- When you place the cursor on a color column, the number of hosts in the range is displayed. You can click a color column to view the list of hosts in the metric range.
  - You can click a host name in the **Host Name** column to access the host details page.
  - Click **View Trends** of a host in the **Operation** column of the list. The metric trends will be displayed. In the current cluster, if you have selected **Host CPU-Memory-Disk Usage**, **View Trends** is unavailable.
- You can click **Export Data** to export the maximum, minimum, and average values of the current metric of all nodes in the cluster within the time range you have specified.

Table 9-4 Metrics

Category	Metric
Process	• Number of Running Processes • Total Number of Processes • Total Number of omm Processes • Total Number of Processes in D and Z States
Network Status	• Host Network Packet Collisions • Number of LAST_ACK States • Number of CLOSING States • Number of LISTENING States • Number of CLOSED States • Number of ESTABLISHED States • Number of SYN_RECV States • Number of TIME_WAITING States • Number of FIN_WAIT2 States • Number of FIN_WAIT1 States • Number of CLOSE_WAIT States • DNS Name Resolution Duration • TCP Ephemeral Port Usage • Host Network Packet Frame Errors

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Metric
Network Reading	• Host Network Read Packets • Host Network Read Dropped Packets • Host Network Read Error Packets • Host Network Receiving Rate
Disk	• Host Disk Write Rate • Host Used Disk • Host Free Disk • Host Disk Read Rate • Host Disk Usage
Memory	• Free Memory • Cache Memory Size • Total Kernel Cache Memory Size • Shared Memory Size • Host Memory Usage • Used Memory
Network Writing	• Host Network Write Packets • Host Network Write Error Packets • Host Network Sending Rate • Host Network Write Dropped Packets
CPU	• CPU Usage of Processes Whose Priorities Have Been Changed • CPU Usage of User Space Processes • CPU Usage of Kernel Space Processes • Host CPU Usage • CPU Total Time • CPU Idle Time
Host Status	• Host File Handle Usage • Average OS Load in 1 Minute • Average OS Load in 5 Minutes • Average OS Load in 15 Minutes • Host PID Usage

Step 3 Click the **Trend** tab to view host monitoring trends.

The resource monitoring trend page of the cluster is displayed. By default, the monitoring data of the past one hour (**1h**) is displayed. You can click to set a time range.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You can click **Add Host to Chart** to add trend lines of up to 12 hosts to the trend charts.
- Click and **Customize** to select the metrics to be displayed on the page.
- You can click **Export Data** to export the maximum, minimum, and average values of all nodes in the cluster for all selected metrics within the time range you have specified.

Step 4 Click the **Cluster** tab to view the cluster information.

You can view the resource monitoring page of each cluster on FusionInsight Manager at the same time.

By default, the monitoring data of the past one hour (**1h**) is displayed. You can click to set a time range.

- Click and **Customize** to select the metrics to be displayed on the page.
- You can click **Export Data** to export the metric values of each cluster within the time range you have specified.

Step 5 Click the **Host** tab to view the host information.

You can view the host resource status, including basic configuration (CPU/memory) and disk configuration.

Click **Export Data** to export the configuration list of all hosts in the cluster.

- Basic Configurations (CPU/Memory)

  You can hover your cursor over the pie chart to view the number of hosts of each hardware configuration in the cluster. The information is displayed in the format of *Number of cores (CPU architecture) Memory size*.

  You can click a slice on the pie chart to view the list of hosts.
- Disk Configurations

  The horizontal axis indicates the total disk capacity (including the OS disk) of a node, and the vertical axis indicates the number of logical disks (including the OS disk).

  You can hover your cursor over a dot to view information about disks of the current configuration, including the quantity of disks, total capacity, and number of hosts.

  You can click a dot on the chart to view the list of hosts.

----End

9.3.4 Dumping MRS Cluster Monitoring Data

The monitoring data reporting function writes the monitoring data collected in the system into a text file and uploads the file to a specified server in FTP or SFTP mode. You can configure interconnection parameters on Manager to save monitoring data to a specified FTP server. In this way, MRS clusters can interconnect with third-party systems.

Manager supports the collection of all the monitoring data in the managed clusters. The collection period is 30 seconds, 60 seconds, or 300 seconds. The monitoring data is stored in different monitoring files on the FTP server by collection period.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- The ECS corresponding to the dump server must be in the same VPC as the Master node of the MRS cluster, and the Master node can access the IP address and a specified port of the dump server.
- The FTP service on the dump server is running properly.

Dumping Monitoring Data of an MRS 3.*x* Cluster

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Interconnection** > **Upload Performance Data**.

Step 3 Toggle on **Upload Performance Data**.

Step 4 Set the upload parameters according to **Table 9-5**.

Table 9-5 Upload parameters

Parameter	Description
FTP IP Address Mode	Specifies the server IP address mode. This parameter is mandatory. The value can be **IPV4** or **IPV6**.
FTP IP Address	Specifies the IP address of the FTP server for storing monitoring files after the monitoring metric data is interconnected. This parameter is mandatory.
FTP Port	Specifies the port for connecting to the FTP server. This parameter is mandatory.
FTP Username	Specifies the username for logging in to the FTP server. This parameter is mandatory.
FTP Password	Specifies the password for logging in to the FTP server. This parameter is mandatory.
Save Path	Specifies the path for storing monitoring files on the FTP server. This parameter is mandatory. The path must be valid.
Dump Interval (second)	Specifies the interval at which monitoring files are periodically stored on the FTP server, in seconds. This parameter is mandatory. The system will periodically upload files to the corresponding FTP server at the specified interval.
Dump Mode	Specifies the protocol used for sending monitoring files. This parameter is mandatory. The value can be **SFTP** or **FTP**. You are advised to use the SFTP mode based on SSH v2. Otherwise, security risks may be incurred.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
SFTP Service Public Key	Specifies the public key of the FTP server. This parameter is optional. It is valid only when **Dump Mode** is set to **SFTP**. **You are advised to configure the parameter to prevent data security risks.**

Step 5 Click **OK**.

After the monitoring metric data dump is configured, the system uploads the files to the corresponding server directory at the specified interval.

Table 9-6 Monitoring data files

Monitoring Data	Description
Collection period	• 30s: real-time metrics that are collected every 30s by default • 60s: real-time metrics that are collected every 60s by default • 300s: all metrics that are not collected every 30s or 60s
File name format	• MRS 3.x and later: **metric**_*Monitoring data collection period*_*File creation time*.**log** Examples: **metric_60_20160908085915.log** and **metric_300_20160908085613.log** • Versions earlier than MRS 3.x: *Cluster name*_**metric**_*Monitoring data collection period*_*File creation time*.**log**
Content format (MRS 3.x as an example)	• Cluster ID|Cluster name|Display name|Service name| **Metric ID**|Collection time|Collection host@m@Sub-metric|Unit|Metric value • Example: 1|xx1|Host|Host|10000413|2019/06/18 10:05:00|10-66-254-146|KB/s|309.910 1|xx1|Host|Host|10000413|2019/06/18 10:05:00|10-66-254-152|KB/s|72.870 2|xx2|Host|Host|10000413|2019/06/18 10:05:00|10-66-254-163|KB/s|100.650

With the **Metric ID** is obtained from the reported file, the third-party can obtain metric details by querying the metric set file in the *Cluster installation path*/**om-server/om/etc/om/all-shown-metric-en_US** on the active/standby OMS node. The file contains the detailed information about all metrics. The content is as follows (taking MRS 3.*x* as an example):

```
Real-Time Metric ID,5-Minute Metric ID,Metric Name,Metric Collection Period (s),Collected by Default,Service Belonged To,Role Belonged To
00101,10000101,JobHistoryServer non-heap memory usage,30,false,Mapreduce,JobHistoryServer
00102,10000102,JobHistoryServer non-heap memory allocation volume,30,false,Mapreduce,JobHistoryServer
00103,10000103,JobHistoryServer heap memory usage,30,false,Mapreduce,JobHistoryServer
00104,10000104,JobHistoryServer heap memory allocation volume,30,false,Mapreduce,JobHistoryServer
00105,10000105,Number of blocked threads,30,false,Mapreduce,JobHistoryServer
00106,10000106,Number of running threads,30,false,Mapreduce,JobHistoryServer
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
00107,10000107,GC time,30,false,Mapreduce,JobHistoryServer
00110,10000110,JobHistoryServer CPU usage,30,false,Mapreduce,JobHistoryServer
...
```

- Metric ID meaning:

  For metrics whose collection period is 30s/60s, you can find the corresponding metric description by referring to the first column, that is, **Real-Time Metric ID**.

  For metrics whose collection period is 300s, you can find the corresponding metric description by referring to the second column, that is, **5-Minute Metric ID**.

- Field description:

  Real-Time Metric ID: indicates the ID of the metric whose collection period is 30s or 60s.

  5-Minute Metric ID: indicates the ID of a 5-minute (300s) metric.

  Metric Collection Period (s): indicates the collection period of real-time metrics. The value can be **30** or **60**.

  Service Belonged To: indicates the name of the service to which a metric belongs, for example, HDFS and HBase.

  Role Belonged To: indicates the name of the role to which a metric belongs, for example, JobServer and RegionServer.

----End

Dumping MRS Cluster Monitoring Data (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In **Configuration**, click **Configure Monitoring Metric Dump** under **Monitoring and Alarm**.

Step 4 Table 9-7 describes the parameters for dumping.

Table 9-7 Parameters

Parameter	Description	Mandatory
Dump Monitoring Metric	Whether to enable monitoring metric data interconnection.	Yes
FTP IP Address	FTP server for storing monitoring files after monitoring metric data is interconnected.	Yes
FTP Port	Port for connecting to the FTP server.	Yes
FTP Username	Username for logging in to the FTP server.	Yes
FTP Password	Password for logging in to the FTP server.	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Mandatory
Save Path	Path for storing monitoring files on the FTP server.	Yes
Dump Interval (second)	Interval at which monitoring files are periodically stored on the FTP server, in seconds. The system will periodically upload files to the corresponding FTP server at the specified interval.	Yes
Dump Mode	Protocol used for sending monitoring files. The options are **FTP** and **SFTP**.	Yes
SFTP Service Public Key	Public key of the FTP server. This parameter is available only when **Dump Mode** is set to **SFTP**. You are advised to configure a public key. Otherwise, security risks may arise.	No

Step 5 Click **OK**.

----End

9.4 Checking MRS Cluster Health

9.4.1 Performing a Health Check for an MRS Cluster

The cluster health check includes three key items: object health, related alarms, and custom metrics. The check result may not match the health status shown on the UI. A health check includes MRS Manager, service-level, and host-level health checks:

- MRS Manager health checks focus on whether the unified management platform can provide management functions properly.
- Service-level health checks focus on whether components can provide services properly.
- Host-level health checks focus on whether host indicators are normal.

To ensure that cluster parameters, configurations, and monitoring are correct and that the cluster can run stably for a long time, you can perform a health check during routine maintenance.

Prerequisites

- The IAM users have been synchronized in advance. You can do this by clicking **Synchronize** next to **IAM User Sync** on the **Dashboard** page of the cluster details.
- You have logged in to MRS Manager. For how to log in, see **Accessing MRS Manager**.

Performing a Health Check on Manager

For MRS 3.*x* and later versions:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Health Check**.

By default, all saved health check reports are displayed in a list. For details, see **Table 9-8**.

Step 3 Start health checks.

- Start a health check.

  Click **Start Check**. In the displayed dialog box, click **OK** to start a health check.

- Set periodic automatic health check.

  Click **Configuration**, select **Enable**, set the check period to **Daily**, **Weekly**, or **Monthly** based on your needs, and click **OK** to save the configuration.

----End

For MRS 2.*x* or earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Start health checks.

- Start a health check.

  Click **Services** and choose **More** > **Start Service Health Check**.

- Set periodic automatic health check.

  a. Click **System**, click **Check Health Status** under **Maintenance**, and click **Configure Health Check**.

  b. Set the **Max. Number of Health Check Reports**. The value must be an integer ranging from 1 to 100.

  c. Click **Periodic Health Check** to enable this function. Set the check period to **Daily**, **Weekly**, or **Monthly** based on O&M requirements, and click **OK** to save the configuration.

----End

Health Check on the MRS Management Console

Only **MRS 1.9.2** clusters support health check management on the MRS console.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Start the health check.

- Perform the health check for all services.

  Choose **Management Operations** > **Start Cluster Health Check**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Perform the health check for a service.

  Click **Components**. In the service list, click the target service name and choose **More** > **Start Service Health Check**.

- Perform the health check for a host.

  Click **Nodes**, expand the node group information, select the check box of the target hosts, and choose **Node** > **Start Host Health Check**.

----End

9.4.2 Performing Health Checks on MRS Cluster Nodes

If the running status of a host is not **Normal**, you can perform health checks on the host to check whether some basic functions are abnormal. During routine O&M, you can perform host health checks to ensure that the configuration parameters and monitoring of each role instance on the host are normal and can run stably for a long time.

Performing Health Checks on Cluster Nodes (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts**.

Step 3 Select the check box of the target host.

Step 4 Select **Health Check** from the **More** drop-down list to start the health check.

To export the result of the health check, click **Export Report** in the upper left corner. If any problem is detected, click **Help**.

----End

Performing Health Checks on Cluster Nodes (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Hosts**.

Step 3 Select the check box of the host for which you want to check the health status.

Step 4 Choose **More** > **Start Host Health Check** to start the health check for the host.

----End

9.4.3 Viewing and Exporting a Health Check Report

You can view the health check result on MRS and export it for further analysis.

Viewing the Health Check Report on the Management Console

Only **MRS 1.9.2** clusters support health check management on the MRS console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, choose **Management Operations** > **View Cluster Health Check Report**.

Step 4 Click **Export Report** on the health check report pane to export the report and view detailed information about check items.

----End

Viewing the Health Check Report on Manager (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Health Check**.

- View the health check report.

  All saved check reports are listed by default. You can filter the check object and result status in the upper right corner of the list. If **Check Type** is **Cluster**, **View Help** is displayed in the **Check Object** drop-down list. During a health check, the system determines whether check objects are healthy based on their historical monitoring metric data.

  Table 9-8 Parameters for a health check report

Item	Description
Check Object	Object to be checked. You can expand the list to view its details.
Status	Check result status. Value options are **No problems found**, **Problems found**, and **Checking**.
Check Type	Entity on which the check is to be performed. Value options are **System**, **Cluster**, **Host**, **Service**, and **OMS**. If you select **Cluster**, all items are checked by default.
Start Mode	Whether the health check is automatically or manually performed
Start Time	Start time of the check
End Time	End time of the check
Operation	Operations you can perform. Value options are **Export Report** and **View Help**.

- Export the health check report.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Locate the row containing the target health check report and click **Export Report** in the **Operation** to download the report.

----End

Viewing the Health Check Report on Manager (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 View or export the health check report.

- View the health check report.
  a. Click **Services**.
  b. Choose **More** > **View Cluster Health Check Report** to view the health check report of a cluster. Click **Export Report** on the health check report pane to export the report and view detailed information about check items.
- Download the health check report.
  a. Choose **System** > **Maintenance** > **Check Health Status**.
  b. Locate the row that contains the target health check report and click **Download** to download the report file.
- Configuring the Number of Health Check Reports to Be Reserved

  Health check reports of MRS clusters, services, and hosts may vary with the time and scenario. You can modify the number of health check reports to be reserved on MRS Manager for later comparison. This setting is valid for health check reports of clusters, services, and hosts. Report files are saved in **$BIGDATA_DATA_HOME/Manager/healthcheck** on the active management node by default and are automatically synchronized to the standby management node.
  a. Choose **System** > **Maintenance** > **Check Health Status** > **Configure Health Check**.
  b. Set **Max. Number of Health Check Reports** to the number of health check reports to be reserved. The value ranges from 1 to 100. The default value is 50.
  c. Click **OK** to save the settings.

----End

9.5 Adjusting the Capacity of an MRS Cluster

9.5.1 Scaling Out an MRS Cluster

The rapid business growth demands a larger storage capacity, a better computing capability, or a higher performance of a cluster. To prevent service interruptions or low efficiency, you can scale out the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can scale out an MRS cluster by adding nodes without changing the system architecture, thereby reducing O&M costs.

Table 9-9 MRS cluster scale-out scenarios

Type	Applicable Scenario	Solution Description
Adding a service	After an MRS cluster is created, you intend to manually add components that were deleted or not installed.	• If the existing resources in the MRS cluster are sufficient for the new service, you can directly **add a service** to the MRS cluster on the console. • If the existing resources in the MRS cluster are insufficient for the new service, add a node group where only basic role instances are deployed in the MRS cluster, and then add services to the node group. For details about how to add a node group, see **Adding a Node Group**.
Scaling out an existing service	You intend to add nodes to an MRS cluster when the resources in the cluster are insufficient for existing components.	• Add nodes to the existing node group in an MRS cluster for scale-out. For details, see **Adding Nodes to a Node Group**. • For task node groups, configure auto scaling rules to implement automatic node scaling. For details, see **MRS Task Node Auto Scaling**.

Prerequisites

- The MRS cluster is in the running state.
- Before scaling out the cluster, ensure that there is an inbound security group rule in which **Protocol & Port** is set to **All**, and **Source** is set to a trusted accessible IP address range.
- You can manually add a node group only for a custom cluster of MRS 3.*x*.

Notes and Constraints

- When you add nodes to a node group where the HBase is installed:

  If automatic DNS registration is not enabled for a node in the cluster, do not start HBase when you expand the node group. Then, update the HBase client configuration by referring to **Updating the MRS Cluster Client After the Server Configuration Expires** and start the HBase instances on the node to be expanded.

  Automatic DNS registration is enabled by default in the following versions:

  MRS 1.9.3, MRS 3.1.0, MRS 3.1.2- LTS, MRS 3.1.5, MRS 3.2.0- LTS, and later versions.

- After a scale-out, the clients installed on nodes in the cluster do not need to be updated. For details about how to update the client installed on nodes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

outside the cluster, see **Updating the MRS Cluster Client After the Server Configuration Expires**.

- The default password of user **root** for the added nodes is the one you set when creating the MRS cluster.
- When you add nodes to a node group where ClickHouse is installed:

  Third- party data source engine tables cannot be automatically synchronized during the scale-out. If such tables are in use, you must manually create them on the newly added nodes after the scale-out. Typical third- party data source engines include MySQL, HDFS, Kafka, and S3.

Adding Nodes to a Node Group

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters**, select a running cluster, and click its name to switch to the cluster details page.

Step 3 Click the **Nodes** tab. In the **Operation** column of the node group, click **Scale Out**. The **Scale Out** page is displayed.

Step 4 Set **System Disk**, **Data Disk**, **scale-out Nodes**, and **Run Bootstrap Action**, and click **OK**.

- **Run Bootstrap Action** parameters may not be supported by clusters of some versions. The operations are subject to the UI.
- If a bootstrap action is added during cluster creation, the **Run Bootstrap Action** parameter is valid. If this function is enabled, the bootstrap actions added during cluster creation will be run on all the scaled out nodes.
- Enabling **Apply New Specifications** parameter indicates that the specifications of the original nodes have been sold out or discontinued. The new nodes will be added based on the new node specifications.

Step 5 In the **Add Node** dialog box, click **OK**.

Step 6 A dialog box is displayed in the upper right corner of the page, indicating that the scale-out task is submitted successfully.

The following parameters explain the cluster scale-out process:

- During scale-out: If a cluster is being scaled out, its status is **Scaling out**. The submitted jobs will be executed and you can submit new jobs. You are not allowed to continue to scale out or delete the cluster. You are advised not to restart the cluster or modify the cluster configuration.
- Successful scale-out: The cluster status is **Running**. The resources used in the old nodes and expanded nodes are charged.
- Failed scale-out: The cluster status is **Running**. You can execute jobs and scale out the cluster again.

After the cluster is scaled out, you can view the node information of the cluster on the **Nodes** page.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Adding a Node Group

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the cluster details page, click the **Nodes** tab and click **Add Node Group**. The **Add Node Group** page is displayed.

Step 4 Set the parameters as required.

Table 9-10 Parameters for adding a node group

Parameter	Description
Name	The node group name must be unique.
CPU Architecture	CPU architecture of an MRS cluster node. The value can be **x86** or **Kunpeng**.
Node Type	• **Task**: A Task node processes data but does not store cluster data. Task nodes support elastic scaling based on service workloads. Data roles that do not store data, such as NodeManager, are deployed on Task nodes. Auto scaling rules can be configured a Task node group. • **Core**: A Core node is an instance that processes and stores data. Component's data storage roles, such as DataNode and Broker, are deployed on Core nodes.
Instance Specifications	Select the instance specifications of the nodes in the current node group. Higher specifications provide stronger data processing and analysis capabilities, but also increase the overall cluster costs. For details about the MRS cluster node specifications, see **MRS Cluster Node Specifications**.
Nodes	Number of nodes in the current node group.
System Disk	Storage type and space of the system disk on a node. You can adjust them as required. For details about the MRS cluster storage, see **Cluster Node Disk Types**.
Data Disk	Storage type and space of data disks on a node. A maximum of 10 disks can be added to each Core or Task node. For details about MRS cluster storage, see **Cluster Node Disk Types**.
Disks	Number of data disks on each node in the node group.
Roles	Component roles to be deployed on nodes in the current node group. Only role instances of components that have been installed in the current cluster can be added.
Billing Mode	Billing mode of the current cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Pricing	Price of the node group to be added.

Step 5 Click **OK**.

After the node group is added, you can view it on the node management page.

----End

Adding a Task Node Group

You can scale out an MRS cluster by manually adding task nodes.

Step 1 Log in to the MRS console.

Step 2 In the **Clusters** area, click the name of the target cluster to go to the cluster details page.

Step 3 On the **Dashboard** page, the cluster type is displayed.

- If the cluster type is **Custom**, go to **Step 4**.
- If the cluster type is not **Custom**, go to **Step 5**.

Step 4 Add Task nodes for a **Custom** cluster.

1. Click the **Nodes** tab then **Add Node Group**. The **Add Node Group** page is displayed.
2. Set **Node Type** to **Task**. Retain the default value **NM** for **Roles** to deploy the NodeManager role. Set other parameters as required.
3. Click **OK** and wait until the node is added. No further action is required.

Step 5 Add Task nodes for a non- custom cluster.

1. Click the **Nodes** tab and then **Configure Task Node**.
2. On the **Configure Task Node** page, set **Node Type**, **Instance Specifications**, **Nodes** and **System Disk**. In addition, if **Add Data Disk** is enabled, configure the storage type, size, and number of data disks.
3. Click **OK** and wait until the node is added.

----End

9.5.2 Scaling In an MRS Cluster

You can remove Core or Task nodes to scale in a cluster based on service requirements. This helps MRS deliver better storage and computing capabilities at lower O&M costs.

The scale-in operation is not allowed for a cluster that is performing active/standby synchronization.

Background

A cluster can contain three types of nodes: Master, Core, and Task nodes. Currently, only Core and Task nodes support scale-in operations. To scale in a

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cluster, you only need to adjust the number of nodes on the MRS console. MRS then automatically selects the nodes to be removed.

The policies for MRS to automatically select nodes are as follows:

- MRS does not select the nodes with basic components installed, such as ZooKeeper, DBService, KrbServer, and LdapServer, because these basic components are the basis for the cluster to run.
- Core nodes store cluster service data. When scaling in a cluster, ensure that all data on the core nodes to be removed has been migrated to other nodes. You can perform follow-up scale-in operations only after all component services are decommissioned, for example, removing nodes from Manager and deleting ECSs. When selecting core nodes, MRS preferentially selects the nodes with a small amount of data and healthy instances to be decommissioned to prevent decommissioning failures. For example, if DataNodes are installed on core nodes in an analysis cluster, MRS preferentially selects the nodes with small data volume and good health status during scale-in.

  When core nodes are removed, their data is migrated to other nodes. If the user business has cached the data storage path, the client will automatically update the path, which may increase the service processing latency temporarily. Cluster scale-in may slow the response of the first access to some HBase on HDFS data. You can restart HBase or disable or enable related tables to resolve this issue.
- Task nodes are computing nodes and do not store cluster data. Data migration is not involved in removing task nodes. Therefore, when selecting task nodes, MRS preferentially selects nodes whose health status is faulty, unknown, or subhealthy. On the **Components** tab of the MRS console, click a service and then the **Instances** tab to view the health status of the node instances.

Scale-In Verification Policy

To prevent component decommissioning failures, components provide different decommissioning constraints. Scale-in is allowed only when the constraints of all installed components are met. **Table 9-11** describes the scale-in verification policies.

Table 9-11 Decommissioning constraints

Component	Constraint
HDFS/DataNode	The number of available nodes after the scale-in is greater than or equal to the number of HDFS copies and the total HDFS data volume does not exceed 80% of the total HDFS cluster capacity. This ensures that the remaining space is sufficient for storing existing data after the scale-in and reserves some space for future use. To ensure data reliability, one backup is automatically generated for each file saved in HDFS, that is, two copies are generated in total.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Constraint
HBase/ RegionServer	The total available memory of RegionServers on all nodes except the nodes to be removed is greater than 1.2 times of the memory which is currently used by RegionServers on these nodes. This ensures that the node to which the region on a decommissioned node is migrated has sufficient memory to bear the region of the decommissioned node.
Kafka/Broker	After the nodes are removed, the number of nodes must be greater than or equal to the maximum number of topic copies and the Kafka disk usage cannot exceed 80% of the Kafka disk space of the entire cluster. (This rule is supported in MRS 2.*x* and earlier versions, as well as in MRS 3.5.0 and later versions.) This prevents insufficient disk space after the scale-in.
Storm/ Supervisor	After the scale-in, ensure that the number of slots in the cluster is sufficient for running the submitted tasks. This prevents no sufficient resources being available for running the stream processing tasks after the scale-in.
Flume/ FlumeServer	If FlumeServer is installed on a node and Flume tasks have been configured for the node, the node cannot be deleted. This prevents the deployed service program from being deleted by mistake.
ClickHouse/ ClickHouseServer	For details, see **Constraints on ClickHouseServer Scale-in**. This ensures that data on the decommissioned nodes is migrated to in-use nodes.
Kudu/ KuduTserver	Rule: When you decommission a Kudu Tserver, all other Kudu instance nodes in the cluster must be normal for the process to succeed. (Only MRS 3.5.0-LTS and later versions support this rule.) Cause: During the decommissioning of a KuduTserver, a rebalance command is executed to migrate the tablets from the decommissioned instance to other KuduTserver nodes in the cluster. If any other KuduTserver nodes are in an abnormal state, the rebalance command will fail, resulting in decommissioning failure.
Yarn/ NodeManager	Rule: NodeManager and ResourceManager are running properly. Reason: NodeManager may be decommissioned unexpectedly, or there are residual containers after decommissioning.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Constraint
Doris/BE (MRS 3.3.0-LTS and later)	Rule: Ensure that at least three BE instances remain after removing the specified instances. Reason: By default, Doris maintains three replicas. If the number of BE instances drops below three, decommissioning will fail and the service may become unavailable.

Scaling In a Cluster by Specifying the Node Quantity

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click the **Nodes** tab. In the **Operation** column of the node group, click **Scale In** to go to the **Scale In** page.

This operation can be performed only when the cluster and all nodes in it are running.

Step 4 Set **Scale-In Type** to **Node quantity**.

Step 5 Set **Scale-In Nodes** and click **OK**.

- Before scaling in the cluster, ensure that there is an inbound security group rule in which **Protocol & Port** is set to **All**, and **Source** is set to a trusted accessible IP address range.
- If there are damaged data blocks in HDFS, the cluster may fail to be scaled in. Contact technical support.

Step 6 A dialog box displayed in the upper right corner of the page indicates that the task of removing the node is submitted successfully.

The cluster scale-in process is explained as follows:

- During scale-in: The cluster status is **Scaling In**. The submitted jobs will be executed, and you can submit new jobs. You are not allowed to continue to scale in or delete the cluster. You are advised not to restart the cluster or modify the cluster configuration.
- Successful scale-in: The cluster status is **Running**.
- Failed scale-in: The cluster status is **Running**. You can execute jobs or scale in the cluster again.

After the cluster is scaled in, you can view the node information of the cluster on the **Nodes** page.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scaling In a Cluster by Removing Nodes that Are No Longer Needed

Delete a node when it is no longer needed. Before deleting a node, decommission the role instance of the component and back up the node's data. For details about how to remove ClickHouseServer nodes, see **Scaling In ClickHouseServer Nodes**.

Step 1 Log in to the MRS console.

Step 2 Click the name of the cluster to go to its details page.

Step 3 Click the **Nodes** tab.

Step 4 Locate the row that contains the target node group and click **Scale In** in the **Operation** column to go to the **Scale In** page.

Step 5 Set **Scale-In Type** to **Specific node** and select the node to be removed.

Nodes in the **Stopped**, **Lost**, **Unknown**, **Isolated**, or **Faulty** status can be specified for scale-in. If the node cannot be selected, click **Stop ECS** to go to the ECS console to stop the node. On the cluster details page of the MRS console, click the **Alarms** tab and check whether any service fault alarms are generated after the node is stopped. If no such an alarm is generated, go back to the **Scale In** page and select the corresponding node for scale-in. If such an alarm is generated, clear the alarm before the scale-in.

Node scale-in may cause data loss. Decommission the node before scale-in. For details, see Decommissioning and Recommissioning an MRS Role Instance.

Step 6 Select **I understand the consequences of performing the scale-in operation**, Click **OK**.

Step 7 Click the **Components** tab and check whether each component is normal. If any component is abnormal, wait for 5 to 10 minutes and check the component status again. If the fault persists, contact technical support.

Step 8 Click the **Alarms** tab and check whether there are exception alarms. If there are exception alarms, clear them before performing other operations.

----End

9.5.3 Scaling In ClickHouseServer Nodes

If ClickHouse is deployed in the MRS cluster, check the data to prevent data loss during node deletion before scaling in ClickHouseServer nodes.

Prerequisites

- This section applies only to MRS 3.1.2-LTS or later.
- Perform ClickHouseServer node scale-in during idle hours because the decommissioning operation occupies bandwidth resources.
- If the logical cluster name of a ClickHouse cluster is changed, contact O&M personnel to update the configuration before decommissioning ClickHouse nodes. Otherwise, the service may become unavailable.
- Before scale-in or decommissioning, ensure that the data on the nodes to be scaled in is either backed up or migrated to other nodes. For instructions on

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

data backup, see **Backing Up ClickHouse Metadata** and **Backing Up ClickHouse Service Data**. For guidance on data migration, contact O&M personnel.

CAUTION

Before migrating data, stop writing data to the nodes that will be decommissioned.

Constraints on ClickHouseServer Scale-in

Table 9-12 Scale-in constraints

Dimension	Constraints
Scale	• If a cluster has only one shard, the instance nodes cannot be removed from the cluster. • Multiple instance nodes in the same shard **must be decommissioned or recommissioned at the same time**. To query cluster shard information, perform the following steps: 1. Log in to the node where the HDFS clients are installed as the client installation user and run the following commands: **cd** *Client installation directory* **source bigdata_env** Security mode **kinit** *ClickHouse service user* **clickhouse client --host** *IP address of the ClickHouse instance* **--port 9440 --secure** In normal mode, run the following command: **clickhouse client --host** *IP address of the ClickHouse instance* **--user** *Username* **--password --port** *9000* *Enter the user password.* 2. Run the following command to query the cluster shard information: **select cluster,shard_num,replica_num,host_name from system.clusters;**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Dimension	Constraints
Cluster Storage	Ensure that the disk space of nodes that will not be decommissioned is sufficient for storing data of all decommissioned nodes. There must be approximately 10% redundant storage space after decommissioning to ensure that the remaining instances can run properly. The procedure is as follows: 1. Run the following command to check the disk usage on each node: **select * from system.disks;** **free_space** indicates the free disk space, and **total_space** indicates the total disk space. The used space is calculated by subtracting the value of **free_space** from that of **total_space**, and its unit is byte. 2. Run the preceding command on a node you want to decommission and calculate the data volume on the node using the preceding formula. 3. Run the preceding command on a node that will not be decommissioned, and then use the following formula: (Value of **free_space** – Data volume of the node to be decommissioned)/Value of **total_space**. If the result is greater than 10%, the node can be decommissioned.
Cluster Status	If there is any faulty ClickHouseServer instance node in the cluster, all instance nodes in the cluster cannot be decommissioned. Log in to Manager, choose **Cluster** > **Services** > **ClickHouse**, click **Instance**, and view the running status of each node in the cluster.
Database	If a database is deployed only on an instance node you want to decommission, the instance node cannot be decommissioned. To remove the instance node, you need to create the database on all ClickHouseServer instance nodes in the cluster. The procedure is as follows: 1. Run the **select * from system.databases;** command to collect the database list of each node. **name** indicates the database name. **engine** indicates the database engine, and the default value is **Atomic**. If the default engine is used, you do not need to specify the engine when creating a table. 2. For the database deployed only on the instance node to be decommissioned, run the following command to create the database: **create database xxx engine=xxx on cluster xxx;**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Dimension	Constraints
Local Non-replicated Table	If a local non-replicated table is deployed only on an instance node you want to decommission, the instance node cannot be decommissioned. To decommission the node, create a local non-replicated table with the same name on any node that will not be decommissioned. For example, the current cluster has two shards, shard 1 has two nodes A and B, and shard 2 has two nodes C and D. The non-replicated table **test** was created without the **ON CLUSTER** keyword, so the table is created only on node A. In this case, to decommission nodes A and B in shard 1, you need to create the table **test** on node C or D in shard 2. Run the following command to list the data tables of each node: **select database,name,engine,create_table_query from system.tables where database != 'system';** Perform the following operations according to the result: ● Check the **engine** column. The table that does not contain the **Replicated** field is a local non-replicated table. ● If there are no replicated tables on any nodes that will not be decommissioned, create one based the table created by **create_table_query**. The following creation statement is an example: **CREATE TABLE {database}.{table} ('column name' type...) ENGINE = MergeTree;**
Replicated Table	If a replicated table exists only on some nodes in the cluster, the nodes where the replicated table is deployed cannot be decommissioned. You need to manually create the replicated table on all instance nodes where no replicated table is deployed in the cluster before decommissioning. For example, the current cluster has two shards, shard 1 has two nodes A and B, and shard 2 has two nodes C and D. The replicated table **test** was created without the **ON CLUSTER** keyword, so the table is created only on nodes A and B. To decommission nodes A and B in shard 1, you need to create the table **test** on nodes C and D in shard 2. Run the following command to list the data tables of each node: **select database,name,engine,create_table_query from system.tables where database != 'system';** Perform the following operations according to the result: ● Check the **engine** column. The table that contains the **Replicated** field is a replicated table. ● If there are no replicated tables on any nodes that will not be decommissioned, create one based the table created by **create_table_query**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Dimension	Constraints
Distributed Table	Distributed tables will not be migrated automatically for decommissioning. Create distributed tables on the nodes that will not be decommissioned. Run the following command to list data tables of each node and check the **engine** column. These tables are distributed tables if this column contains field **Distributed**. **select database,name,engine from system.tables where database != 'system';** Creating distributed tables on these nodes will not affect the decommissioning, but may affect subsequent service operations.
View	Views will not be automatically migrated for decommissioning, and views do not store data. Run the following command to list data tables of each node and check the **engine** column. These tables are views if this column contains field **View**. **select database,name,engine from system.tables where database != 'system';** Run the following command to delete the views one by one: **drop view {database_name}.{table_name};**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Dimension	Constraints
Materialized Views	Materialized views will not be automatically migrated for Decommissioning. Create materialized views on the nodes that will not be decommissioned. If the materialized view of a node to be decommissioned does not display the specified aggregation table but uses an embedded table, the node cannot be decommissioned. Run the following command to list data tables of each node and check the **engine** column. These tables are materialized views if this column contains field **MaterializedView**. **select database,name,engine, create_table_query from system.tables where database != 'system';** Embedded tables are initialized with a **POPULATE** field in their **create_table_query** column. Views are created at initialization, and new data is ignored. Aggregation tables without the **POPULATE** field insert new data directly into views and support tables, requiring manual data loading. The table creation operations of the aggregation table and embedded table are different. Perform the following operations to process the materialized views of the node to be decommissioned: 1. Record the materialized views and delete them. **drop view {database_name}.{table_name};** 2. After the node decommissioning is complete, delete and recreate the corresponding materialized views on in-use nodes to update the materialized views. 3. To create an aggregation table, specify **WHERE** to search for historical data and manually import the historical data to the materialized views. Otherwise, historical data cannot be imported to the materialized views based on unified conditions. As a result, data is imported repeatedly. For example, an update point can be specified to ensure that data before the update point is manually loaded in **INSERT** mode. ● Add **WHERE** *{ Time field (for example, date)}***>= toDate (***{ Current time (for example, '2022-12-01 00:00:00')}***)** to the table creation statement. ● **insert into** *{table}* **select** *{Table field}* **from** *{Source table}* **where** *{Time field}***< toDate** *({Current time})* is used to load original data. 4. Embedded tables will lose data generated during table creation. You can specify **WHERE** to filter out all historical data. In this case, an empty table is created, and you only need to manually insert all data in the historical data source table.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Dimension	Constraints
Tables of Third-Party Engines	Currently, tables of third-party engines cannot be automatically migrated for decommissioning. Run the following command to list data tables of each node and check the **engine** column. These tables are tables of third-party engines if this column does not contain any of the following fields: **MergeTree**, **View**, **MaterializedView**, **Distributed**, and **Log**. (The **engine** column of a third-party engine table may contain field **Memory**, **HDFS**, or **MySQL**.) **select database,name,engine from system.tables where database != 'system';** Create third-party engine tables on the nodes that will not be decommissioned and delete those from the nodes that will be decommissioned.
Detached Data	If the table on a node to be decommissioned has been detached and data still exists in the **detached** directory, the node cannot be decommissioned. You need to attach the data in the **detached** directory to other directories before decommissioning. 1. Run the following command to view the **system.detached_parts** system catalog of the node to be decommissioned: **select * from system.detached_parts;** 2. If **detached part** data exists and these partitions are no longer used, run the following command to delete the **detached part** data: **ALTER TABLE {table_name} DROP DETACHED PARTITION {partition_expr} SETTINGS allow_drop_detached = 1;** 3. Run the following command to check whether there is any **detached part** data in the **system.detached_parts** system catalog: **select * from system.detached_parts;** If the command output is empty, there is no **detached part** data in this system catalog.

Scaling In ClickHouseServer Nodes

Before removing ClickHouseServer instance nodes, you need to decommission them. Multiple node replicas of the same shard **must be decommissioned at the same time**. If there is a faulty ClickHouseServer instance node in the cluster, all instance nodes of the cluster cannot be decommissioned. For more constraints, see **Constraints on ClickHouseServer Scale-in**.

Decommissioning can be performed only on ClickHouseServer nodes. ClickHouseBalancer nodes cannot be decommissioned.

Step 1 Log in to FusionInsight Manager of the MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > **ClickHouse** > **Instances**. Click the ClickHouseServer instance to be decommissioned. On the displayed **Dashboard** page, click **metrika.xml** in the **Configuration File** area.

Modify the **<shard>*</shard>** and **<host>*</host>** parts.

- **<shard>*</shard>** indicates the shard ID of the node.
- **<host>*</host>** indicates the node name. Two nodes in the same shard act as replicas of each other.

```
<yandex>
 <macros>
  <id>90</id>
  <shard>89</shard>
  <replica>1</replica>
 </macros>
 <ZK_ROOT_PATH>/clickhouse</ZK_ROOT_PATH>
 <clickhouse_remote_servers>
  <default_cluster>
   <shard>
    <internal_replication>true</internal_replication>
    <replica>
     <host>server-2110081635-0001</host>
     <port>9000</port>
     <user>clickhouse</user>
     <password/>
    </replica>
    <replica>
     <host>server-2110081635-0002</host>
     <port>9000</port>
     <user>clickhouse</user>
     <password/>
    </replica>
   </shard>
  </default_cluster>
 </clickhouse_remote_servers>
</yandex>
```

Step 3 Log in to the MRS console and click the cluster name to go to the cluster details page.

Step 4 Click the **Components** tab and click **ClickHouse**. Then switch to **Instances**, select the **ClickHouseServer** instances to be removed, click **More**, and select **Decommission**.

Step 5 Click the **Components** tab and click **ClickHouse**. Then click **More**, and select **Synchronize Configuration**.

Step 6 Click the **Nodes** tab and click the ClickHouseServer instance node that has been decommissioned.

Step 7 On the ECS page, click **Stop**. In the displayed dialog box, select **Forcibly stop the preceding ECSs** and click **Yes**.

Step 8 Go back to the MRS console, click the **Nodes** tab, locate the row that contains the target node group, and click **Scale In** in the **Operation** column to go to the **Scale In** page.

Step 9 Set **Scale-In Type** to **Specific node** and select the node to be removed.

Step 10 Select **I understand the consequences of performing the scale-in operation**. Click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 Click the **Components** tab and check whether each component is normal. If any component is abnormal, wait for 5 to 10 minutes and check the component status again. If the fault persists, contact technical support.

Step 12 Click the **Alarms** tab and check whether there are exception alarms. If there are exception alarms, clear them before performing other operations.

----End

9.5.4 MRS Task Node Auto Scaling

9.5.4.1 Automatic Scaling of Task Nodes in an MRS Cluster

In big data application scenarios, especially real-time data analysis and processing, the number of cluster nodes needs to be dynamically adjusted according to data volume changes to ensure that the required resources are provided.

MRS supports auto scaling of Task nodes based on cluster loads. If the data volume changes periodically, you can configure an auto scaling rule so that the number of task nodes can be automatically adjusted in a fixed period of time before the data volume changes.

- Load-based auto scaling of Task nodes: The number of Task nodes is adjusted based on the real-time cluster load metrics or cluster resource pool metrics. When the data volume changes, scaling is triggered with a delay. For details about the metrics that can be configured, see **Node Auto Scaling Metrics**.
- Time-based Task node auto scaling: The number of Task nodes is adjusted based on the time ranges. If the data volume changes periodically, you can specify a resource plan to scale in or out the cluster before the data volume changes so as to avoid scaling delay.
- Automation script: In some scaling scenarios, you may need to manually reallocate resources or modify service logic based on the number of nodes after cluster scaling. MRS allows you to customize automation scripts for auto scaling to solve this problem. Automation scripts can be executed to adjust clusters to service load changes, freeing you from manual operations. In addition, the automation scripts can be fully customized according to various requirements, making the auto scaling more flexible. For details about how to configure automation scripts, see **Configuring Bootstrap Actions for an MRS Cluster Node**.

You can configure load-based or time-based auto scaling of Task nodes. You can also configure both of them, which is recommended in case of unexpected data peaks.

Node Auto Scaling Metrics

- **Node group dimension policy**

  When adding a rule, you can refer to **Table 9-13** to configure the corresponding metrics.

  – When the value type is percentage or ratio in **Table 9-13**, the valid value can be accurate to percentile. The percentage-type metric values omit the percent sign (%). For example, 16.80 indicates 16.80%.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

– Hybrid clusters support all metrics of analytics clusters and streaming clusters.

Table 9-13 Auto scaling metrics

Cluster Type	Metric	Value Type	Description
Streaming cluster	StormSlotAvailable	Integer	Number of available Storm slots Value range: 0 to 2147483646
	StormSlotAvailablePercentage	Percentage	Percentage of available Storm slots, that is, the proportion of the available slots Value range: 0 to 100
	StormSlotUsed	Integer	Number of the used Storm slots Value range: 0 to 2147483646
	StormSlotUsedPercentage	Percentage	Percentage of the used Storm slots, that is, the proportion of the used slots Value range: 0 to 100
	StormSupervisorMemAverageUsage	Integer	Average memory usage of the Supervisor process of Storm Value range: 0 to 2147483646
	StormSupervisorMemAverageUsagePercentage	Percentage	Average percentage of the memory used by the Supervisor process of Storm to the total memory of the system Value range: 0 to 100
	StormSupervisorCPUAverageUsagePercentage	Percentage	Average percentage of the CPUs used by the Supervisor process of Storm to the total CPUs Value range: 0 to 6000
Analysis cluster	YARNAppPending	Integer	Number of pending tasks on YARN Value range: 0 to 2147483646
	YARNAppPendingRatio	Ratio	Ratio of pending tasks on YARN, that is, the ratio of pending tasks to running tasks on YARN Value range: 0 to 2147483646
	YARNAppRunning	Integer	Number of running tasks on YARN Value range: 0 to 2147483646

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Cluster Type	Metric	Value Type	Description
	YARNContainerAllocated	Integer	Number of containers allocated to YARN Value range: 0 to 2147483646
	YARNContainerPending	Integer	Number of pending containers on YARN Value range: 0 to 2147483646
	YARNContainerPendingRatio	Ratio	Ratio of pending containers on YARN, that is, the ratio of pending containers to running containers on YARN. Value range: 0 to 2147483646
	YARNCPUAllocated	Integer	Number of virtual CPUs (vCPUs) allocated to YARN Value range: 0 to 2147483646
	YARNCPUAvailable	Integer	Number of available vCPUs on YARN Value range: 0 to 2147483646
	YARNCPUAvailablePercentage	Percentage	Percentage of available vCPUs on YARN, that is, the proportion of available vCPUs to total vCPUs Value range: 0 to 100
	YARNCPUPending	Integer	Number of pending vCPUs on YARN Value range: 0 to 2147483646
	YARNMemoryAllocated	Integer	Memory allocated to YARN. The unit is MB. Value range: 0 to 2147483646
	YARNMemoryAvailable	Integer	Available memory on YARN. The unit is MB. Value range: 0 to 2147483646
	YARNMemoryAvailablePercentage	Percentage	Percentage of available memory on YARN, that is, the proportion of available memory to total memory on YARN Value range: 0 to 100
	YARNMemoryPending	Integer	Pending memory on YARN Value range: 0 to 2147483646

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Resource pool policy**

  Only the following versions support auto scaling by resource pool:

  – Common MRS: MRS 3.1.5 or later

  – LTS MRS: MRS 3.3.0-LTS or later

  When adding a rule, you can refer to **Table 9-14** to configure the corresponding metrics.

  Table 9-14 Rule configuration description

Cluster Type	Metric	Value Type	Description
Analysis/Custom cluster	ResourcePoolMemoryAvailable	Integer	Available memory on YARN in the resource pool. The unit is MB. Value range: 0 to 2147483646
	ResourcePoolMemoryAvailablePercentage	Percentage	Percentage of available memory on YARN in the resource pool, that is, the proportion of available memory to total memory on YARN Value range: 0 to 100
	ResourcePoolCPUAvailable	Integer	Number of available vCPUs on YARN in the resource pool Value range: 0 to 2147483646
	ResourcePoolCPUAvailablePercentage	Percentage	Percentage of available vCPUs on YARN in the resource pool. that is, the proportion of available vCPUs to total vCPUs Value range: 0 to 100

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

When adding a resource plan, you can set parameters by referring to **Table 9-15**.

Table 9-15 Configuration items of a resource plan

Configuration Item	Example Value	Description
Effective On	**Monday**	The effective date of a resource plan. **Daily** is selected by default. You can also select one or multiple days from Monday to Sunday.
Time Range	**08:00-10:00**	Start time and End time of a resource plan are accurate to minutes, with the value ranging from **00:00** to **23:59**. For example, if a resource plan starts at 8:00 and ends at 10:00, set this parameter to 8:00-10:00. The end time must be at least 30 minutes later than the start time.
Node Range	**4-5**	The number of nodes in a resource plan ranges from **0** to **500**. In the time range specified in the resource plan, if the number of Task nodes is less than the specified minimum number of nodes, it will be increased to the specified minimum value of the node range at a time. If the number of Task nodes is greater than the maximum number of nodes specified in the resource plan, the auto scaling function reduces the number of Task nodes to the maximum value of the node range at a time. The minimum number of nodes must be less than or equal to the maximum number of nodes.

- When a resource plan is enabled, the **Default Range** value on the auto scaling page forcibly takes effect beyond the time range specified in the resource plan. For example, if **Default Range** is set to **1-2**, **Time Range** is between **08:00-10:00**, and **Node Range** is **4-5** in a resource plan, the number of Task nodes in other periods (0:00-8:00 and 10:00-23:59) of a day is forcibly limited to the default node range (1 to 2). If the number of nodes is greater than 2, auto scale-in is triggered; if the number of nodes is less than 1, auto scale-out is triggered.
- When a resource plan is not enabled, the **Default Range** takes effect in all time ranges. If the number of nodes is not within the default node range, the number of Task nodes is automatically increased or decreased to the default node range.
- Time ranges of resource plans cannot be overlapped. The overlapped time range indicates that two effective resource plans exist at a time point. For example, if resource plan 1 takes effect from **08:00** to **10:00** and resource plan 2 takes effect from **09:00** to **11:00**, the time range between **09:00** to **10:00** is overlapped.
- The time range of a resource plan must be on the same day. For example, you cannot configure a resource plan from **23:00** to **01:00** (the next day).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Instead, you need to configure two resource plans, one from 23:00 to 00:00 and the other from 00:00 to 01:00.

- **Automation scripts**

  When adding an automation script, you can set related parameters by referring to **Table 9-16**.

  Table 9-16 Configuration items of an automation script

Configuration Item	Example Value	Description
Name	test	Automation script name. • The value can contain only digits, letters, spaces, hyphens (-), and underscores (_) and cannot start with a space. • The value can contain 1 to 64 characters. • A name must be unique in the same cluster. You can set the same name for different clusters.
Script Path	obs://mrs-samples/test.sh	Script path. The value can be an OBS file system path or a local VM path. • An OBS file system path must start with **obs://** and end with **.sh**, for example, **obs://mrs-samples/*xxx*.sh**. • A local VM path must start with a slash (/) and end with **.sh**. For example, the path of the example script for installing the Zepelin is **/opt/bootstrap/zepelin/zepelin_install.sh**.
Execution Node	Master	Select a type of the node where an automation script is executed. If you select **Master** nodes, you can choose whether to run the script only on the active Master nodes by enabling or disabling **Active Master**. • If you enable it, the script runs only on the active Master nodes. • If you disable it, the script runs on all Master nodes. It is disabled by default.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configuration Item	Example Value	Description
Parameter	-	Automation script parameter. The following predefined variables can be imported to obtain auto scaling information: • **${mrs_scale_node_num}**: Number of auto scaling nodes. The value is always positive. • **${mrs_scale_type}**: Scale-out/in type. The value can be **scale_out** or **scale_in**. • **${mrs_scale_node_hostnames}**: Host names of the auto scaling nodes. Use commas (,) to separate multiple host names. • **${mrs_scale_node_ips}**: IP address of the auto scaling nodes. Use commas (,) to separate multiple IP addresses. • **${mrs_scale_rule_name}**: Name of the triggered auto scaling rule. For a resource plan, this parameter is set to **resource_plan**.
Executed	Before scale-out	Time for executing an automation script. The following four options are supported: **Before scale-out**, **After scale-out**, **Before scale-in**, and **After scale-in**. Assume that the execution nodes include Task nodes. • The automation script executed before scale-out cannot run on the Task nodes to be added. • The automation script executed after scale-out can run on the added Task nodes. • The automation script executed before scale-in can run on the Task nodes to be deleted. • The automation script executed after scale-in cannot run on the deleted Task nodes.
Action upon Failure	Continue	Whether to continue to execute subsequent scripts and scale-out/in after the script fails to be executed. • You are advised to set this parameter to **Continue** in the commissioning phase so that the cluster can continue the scale-out/in operation no matter whether the script is executed successfully. • If the script fails to be executed, view the failure logs in the **/var/log/Bigdata/Bootstrap** directory on the cluster VM. • The scale-in operation cannot be rolled back. Therefore, the **Action upon Failure** can only be set to **Continue** after scale-in.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The automation script is triggered only during auto scaling. It is not triggered when the cluster node is manually scaled out or in.

9.5.4.2 Adding an Auto Scaling Policy for MRS Task Nodes

The auto scaling function of MRS enables the task nodes of a cluster to be automatically scaled to match cluster loads. If the data volume changes periodically, you can configure an auto scaling rule so that the number of task nodes can be automatically adjusted in a fixed period of time before the data volume changes.

- Load-based auto scaling of Task nodes: The number of Task nodes is adjusted based on the real-time cluster load metrics or cluster resource pool metrics. When the data volume changes, scaling is triggered with a delay.
- Time-based Task node auto scaling: The number of Task nodes is adjusted based on the time ranges. If the data volume changes periodically, you can specify a resource plan to scale in or out the cluster before the data volume changes to avoid a scaling delay.

You can configure load-based or time-based auto scaling of Task nodes. You can also configure both of them, which is recommended in case of unexpected data peaks.

For more information about auto scaling policies, see **Automatic Scaling of Task Nodes in an MRS Cluster**.

Configuring Auto Scaling During Cluster Creation

When you create an **analysis, streaming, or hybrid** cluster, you can configure auto scaling policies in advanced configuration parameters.

Step 1 Log in to the MRS console.

Step 2 When you create a cluster containing task nodes, configure the cluster software and hardware information by referring to **Creating a Custom MRS Cluster**. Then, on the **Set Advanced Options** page, enable **Analysis Task** and configure or modify auto scaling rules and resource plans.

You can configure auto scaling in the following scenarios.

----End

Scenario 1: Configuring Real-Time Load-based Auto Scaling

When it comes to adjusting the number of Task nodes based on the real-time cluster load or cluster resource pool metrics, you can use the MRS auto scaling rules.

Take the YARN service as an example. If the ratio of available YARN memory is less than 20%, five nodes will be added. If the ratio of available YARN memory is greater than 70%, five nodes will be reduced. The number of nodes in a task node group ranges from 1 to 10.

Step 1 Go to the **Auto Scaling** page to configure auto scaling rules.

- Configure the **Default Range** parameter.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Enter a task node range, in which auto scaling is performed. This constraint applies to all scale-in and scale-out rules. The maximum value range allowed is 0 to 500.

The value range in this example is 1 to 10.

- Configure an auto scaling rule.

  To enable **Auto Scaling**, you must configure a scale-out or scale-in rule.

  a. Select **Scale-Out** or **Scale-In**.

  b. Click **Add Rule**.

     - You can set a maximum of five rules for scaling out or in a cluster, respectively.
     - The system determines the scale-out and then scale-in based on your configuration sequence. Add important policies first. This prevents repeated scale-out or scale-in operations in case a scale-out or scale-in cannot meet requirements.

  c. Set parameters as required.

Table 9-17 Parameters for adding a scaling rule

Parameter	Description	Example Value
Rule Name	Name of an auto scaling rule	default-expand-1
If	The metric threshold of an auto scaling rule. ▪ For details about auto scaling metrics, see **Table 9-13** in **Automatic Scaling of Task Nodes in an MRS Cluster**. ▪ Comparison factors include greater than, greater than or equal to, less than, and less than or equal to.	YARNMemoryAvailablePercentage Less than or equal to 20%
Last For	The duration (5n) during which a cluster reaches the configured metric threshold for triggering a scale-out or scale-in. The default value of **n** is 1, indicating that the duration is five minutes.	1
Add	Number of Task nodes to be added or removed at a time. The value ranges from 1 to 100.	5
Cooldown Period	Interval between two auto scaling operations. The default cooldown period is 20 minutes. The minimum value is **0**.	20

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

d. Click **OK**.

You can view, edit, or delete the rules you configured in the **Scale-out** or **Scale-in** area on the **Auto Scaling** page. You can click **Add Rule** to configure multiple rules.

Step 2 Click **OK**.

- If you are configuring an auto scaling rule for an existing cluster, select **I agree to authorize MRS to scale out or in nodes based on the above rule**.
- After the configuration, system automatically scales in or out the Task node group as long as the auto scaling rules are met.
- You can click **Operation Logs** in the navigation pane to view the scaling details of a cluster.

----End

Scenario 2: Configuring Time-based Auto Scaling

If the data volume changes regularly every day and you want to scale out or in a cluster before the data volume changes, you can create resource plans to adjust the number of Task nodes as planned in the specified time range.

For example: A real-time processing service sees a sharp increase in data volume from 7:00 to 13:00 on Monday, Tuesday, and Saturday. Assume that an MRS streaming cluster is used to process the service data. Five task nodes are required from 7:00 to 13:00 on Monday, Tuesday, and Saturday, while only two are required at other time.

Step 1 Go to the **Auto Scaling** page to configure a resource plan.

Step 2 For example, the **Default Range** of node quantity is set to **2-2**, indicating that the number of task nodes is fixed at 2 except the time range specified in the resource plan.

Step 3 Click **Configure Node Range for Specific Time Range** under **Default Range** or **Add Resource Plan**.

Step 4 Configure **Effective On**, **Time Range**, and **Node Range**.

You can click **Configure Node Range for Specific Time Range** to configure multiple resource plans. A maximum of five resource plans are supported.

After a resource plan triggers a scale-out or scale-in, auto scaling will not be triggered again within 10 minutes.

Table 9-18 Scaling rule parameters

Parameter	Description	Example Value
Effective On	**Effective On** is set to **Daily** by default. You can also select one or more days from Monday to Sunday.	Monday, Tuesday, and Saturday

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Time Range	The start time must be at least 30 minutes earlier than the end time. Time ranges configured for different resource plans cannot overlap.	07:00-13:00
Node Range	• If you set both **Node Range** and **Time Range**, the node range you set will be used during the time range you set, and the default node range will be used beyond the time range you set. • If you do not set **Node Range**, its default value will be used.	5-5

Step 5 Click **OK**.

- If you are configuring a resource plan for an existing cluster, select **I agree to authorize MRS to scale out or in nodes based on the above rule**.
- After the configuration, system automatically scales in or out the Task node group as long as the auto scaling rules are met.
- You can click **Operation Logs** in the navigation pane to view the scaling details of a cluster.

----End

Scenario 3: Configuring Time-based and Load-based Auto Scaling

If the data volume is not stable and the expected fluctuation may occur, the fixed Task node range cannot guarantee that the requirements in some service scenarios are met. In this case, it is necessary to adjust the number of Task nodes based on the real-time loads and resource plans.

For example: A real-time processing service sees an unstable increase in data volume from 7:00 to 13:00 on Monday, Tuesday, and Saturday. In these time ranges, five to eight Task nodes are required. In other time ranges, two to four Task nodes are required. You can set an auto scaling rule based on a resource plan. When the data volume exceeds the expected value, the number of Task nodes changes with resource loads, without exceeding the node range specified in the resource plan. When a resource plan is triggered, the number of nodes is adjusted within the specified node range with minimum affect. That is, increase nodes to the upper limit and decrease nodes to the lower limit.

Step 1 Go to the **Auto Scaling** page to configure auto scaling rules.

- Default Range

  Enter a task node range, in which auto scaling is performed. This constraint applies to all scale-in and scale-out rules.

  The number of task nodes in a cluster is limited to the configured default node range or the node range in the resource plan that takes effect in the current time range. The node quantity range in the resource plan has a higher priority.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, this parameter is set to **2-4** in this scenario.

- Configuring an Auto Scaling Rule

  To enable **Auto Scaling**, you must configure a scale-out or scale-in rule.

  a. Select **Scale-Out** or **Scale-In**.

  b. Click **Add Rule**. The **Add Rule** page is displayed.

     - You can set a maximum of five rules for scaling out or in a cluster, respectively.
     - The system determines the scale-out and then scale-in based on your configuration sequence. Add important policies first. This prevents repeated scale-out or scale-in operations in case a scale-out or scale-in cannot meet requirements.

  c. Set parameters as required.

     Table 9-19 Parameters for adding a rule

Parameter	Description	Example Value
Rule Name	Name of an auto scaling rule	default-expand-2
If	The metric threshold of an auto scaling rule. ▪ For details about auto scaling metrics, see **Table 9-13** in **Automatic Scaling of Task Nodes in an MRS Cluster**. ▪ Comparison factors include greater than, greater than or equal to, less than, and less than or equal to.	YARNAppPending Greater than 75
Last For	The duration (5n) during which a cluster reaches the configured metric threshold for triggering a scale-out or scale-in. The default value of **n** is 1, indicating that the duration is five minutes.	1
Add	Number of Task nodes to be added or removed at a time. The value ranges from 1 to 100.	1
Cooldown Period	Interval between two auto scaling operations. The default cooldown period is 20 minutes. The minimum value is **0**.	20

  d. Click **OK**.

     You can view, edit, or delete the rules you configured in the **Scale-out** or **Scale-in** area on the **Auto Scaling** page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Configure a resource plan.

1. Click **Configure Node Range for Specific Time Range** under **Default Range** or **Add Resource Plan**.
2. Configure **Effective On**, **Time Range**, and **Node Range**.

   Click **Configure Node Range for Specific Time Range** or **Add Resource Plan** to configure multiple resource plans.

   After a resource plan triggers a scale-out or scale-in, auto scaling will not be triggered again within 10 minutes.

Table 9-20 Scaling rule parameters

Parameter	Description	Example Value
Effective On	**Effective On** is set to **Daily** by default. You can also select one or more days from Monday to Sunday.	Monday, Tuesday, and Saturday
Time Range	The start time must be at least 30 minutes earlier than the end time. Time ranges configured for different resource plans cannot overlap.	07:00-13:00
Node Range	– If you set both **Node Range** and **Time Range**, the node range you set will be used during the time range you set, and the default node range will be used beyond the time range you set. – If you do not set **Node Range**, its default value will be used.	5-8

Step 3 Click **OK**.

- If you are configuring a resource plan for an existing cluster, select **I agree to authorize MRS to scale out or in nodes based on the above rule**.
- After the configuration, system automatically scales in or out the Task node group as long as the auto scaling rules are met.
- You can click **Operation Logs** in the navigation pane to view the scaling details of a cluster.

----End

Adding an Auto Scaling Policy for a Cluster

After a cluster is created, you can configure rules for the task node group in a cluster by node group or resource pool.

The node group policy and resource pool policy are mutually exclusive. You can configure either of them as needed.

You can specify auto scaling policies by resource pool in MRS 3.1.5 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Item	By Node Group	By Resource Pool
Auto scaling object	All nodes in the task node group	Task nodes in the resource pool specified by an auto scaling policy
Resource pool ownership of added nodes	Default resource pool	Resource pool specified by the auto scaling policy
Scale-in object	Random scale-in of nodes in the task node group	Random scale-in of nodes in a resource pool specified by an auto scaling policy

Notes and Constraints

- Auto scaling policies can be enabled for only one node group.
- Only the following versions support auto scaling by resource pool:
  - Common MRS: MRS 3.1.5 or later
  - LTS MRS: MRS 3.3.0-LTS or later

Prerequisites

- A task node group has been configured by referring to **Adding a Task Node**.
- A resource pool has been added when resource pool-based auto scaling is required. For details, see **Adding an MRS Tenant Resource Pool**.

Procedure

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, click the name of the target cluster. The cluster details page is displayed.

Step 3 On the page that is displayed, click the **Auto Scaling** tab.

You can configure policies by resource pool or node group as needed.

Step 4 Click **Create Auto Scaling Policy** to create an auto scaling policy.

You can configure auto scaling in the following scenarios.

----End

Scenario 1: Configuring Real-Time Load-based Auto Scaling

When it comes to adjusting the number of Task nodes based on the real-time cluster load or cluster resource pool metrics, you can use the MRS auto scaling rules.

Take the YARN service as an example. If the ratio of available YARN memory is less than 20%, five nodes will be added. If the ratio of available YARN memory is

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

greater than 70%, five nodes will be reduced. The number of nodes in a task node group ranges from 1 to 10.

Step 1 Go to the **Auto Scaling** page to configure auto scaling rules.

- Configure the **Default Range** parameter.

  Enter a task node range, in which auto scaling is performed. This constraint applies to all scale-in and scale-out rules. The maximum value range allowed is 0 to 500.

  The value range in this example is 1 to 10.

- Configure an auto scaling rule.

  To enable **Auto Scaling**, you must configure a scale-out or scale-in rule.

  a. Select **Scale-Out** or **Scale-In**.

  b. Click **Add Rule**.

  - You can set a maximum of five rules for scaling out or in a cluster, respectively.
  - The system determines the scale-out and then scale-in based on your configuration sequence. Add important policies first. This prevents repeated scale-out or scale-in operations in case a scale-out or scale-in cannot meet requirements.

  c. Set parameters as required.

Table 9-21 Parameters for adding a scaling rule

Parameter	Description	Example Value
Rule Name	Name of an auto scaling rule	default-expand-1
If	The metric threshold of an auto scaling rule. ■ For details about auto scaling metrics, see **Table 9-13** in **Automatic Scaling of Task Nodes in an MRS Cluster**. ■ Comparison factors include greater than, greater than or equal to, less than, and less than or equal to.	YARNMemoryAvailablePercentage Less than or equal to 20%
Last For	The duration (5n) during which a cluster reaches the configured metric threshold for triggering a scale-out or scale-in. The default value of **n** is 1, indicating that the duration is five minutes.	1
Add	Number of Task nodes to be added or removed at a time. The value ranges from 1 to 100.	5

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Cooldown Period	Interval between two auto scaling operations. The default cooldown period is 20 minutes. The minimum value is **0**.	20

d. Click **OK**.

You can view, edit, or delete the rules you configured in the **Scale-out** or **Scale-in** area on the **Auto Scaling** page. You can click **Add Rule** to configure multiple rules.

Step 2 Click **OK**.

- If you are configuring an auto scaling rule for an existing cluster, select **I agree to authorize MRS to scale out or in nodes based on the above rule**.
- After the configuration, system automatically scales in or out the Task node group as long as the auto scaling rules are met.
- You can click **Operation Logs** in the navigation pane to view the scaling details of a cluster.

----End

Scenario 2: Configuring Time-based Auto Scaling

If the data volume changes regularly every day and you want to scale out or in a cluster before the data volume changes, you can create resource plans to adjust the number of Task nodes as planned in the specified time range.

For example: A real-time processing service sees a sharp increase in data volume from 7:00 to 13:00 on Monday, Tuesday, and Saturday. Assume that an MRS streaming cluster is used to process the service data. Five task nodes are required from 7:00 to 13:00 on Monday, Tuesday, and Saturday, while only two are required at other time.

Step 1 Go to the **Auto Scaling** page to configure a resource plan.

Step 2 For example, the **Default Range** of node quantity is set to **2-2**, indicating that the number of task nodes is fixed at 2 except the time range specified in the resource plan.

Step 3 Click **Configure Node Range for Specific Time Range** under **Default Range** or **Add Resource Plan**.

Step 4 Configure **Effective On**, **Time Range**, and **Node Range**.

You can click **Configure Node Range for Specific Time Range** to configure multiple resource plans. A maximum of five resource plans are supported.

After a resource plan triggers a scale-out or scale-in, auto scaling will not be triggered again within 10 minutes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-22 Scaling rule parameters

Parameter	Description	Example Value
Effective On	**Effective On** is set to **Daily** by default. You can also select one or more days from Monday to Sunday.	Monday, Tuesday, and Saturday
Time Range	The start time must be at least 30 minutes earlier than the end time. Time ranges configured for different resource plans cannot overlap.	07:00-13:00
Node Range	• If you set both **Node Range** and **Time Range**, the node range you set will be used during the time range you set, and the default node range will be used beyond the time range you set. • If you do not set **Node Range**, its default value will be used.	5-5

Step 5 Click **OK**.

- If you are configuring a resource plan for an existing cluster, select **I agree to authorize MRS to scale out or in nodes based on the above rule**.
- After the configuration, system automatically scales in or out the Task node group as long as the auto scaling rules are met.
- You can click **Operation Logs** in the navigation pane to view the scaling details of a cluster.

----End

Scenario 3: Configuring Time-based and Load-based Auto Scaling

If the data volume is not stable and the expected fluctuation may occur, the fixed Task node range cannot guarantee that the requirements in some service scenarios are met. In this case, it is necessary to adjust the number of Task nodes based on the real-time loads and resource plans.

For example: A real-time processing service sees an unstable increase in data volume from 7:00 to 13:00 on Monday, Tuesday, and Saturday. In these time ranges, five to eight Task nodes are required. In other time ranges, two to four Task nodes are required. You can set an auto scaling rule based on a resource plan. When the data volume exceeds the expected value, the number of Task nodes changes with resource loads, without exceeding the node range specified in the resource plan. When a resource plan is triggered, the number of nodes is adjusted within the specified node range with minimum affect. That is, increase nodes to the upper limit and decrease nodes to the lower limit.

Step 1 Go to the **Auto Scaling** page to configure auto scaling rules.

- Default Range

  Enter a task node range, in which auto scaling is performed. This constraint applies to all scale-in and scale-out rules.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The number of task nodes in a cluster is limited to the configured default node range or the node range in the resource plan that takes effect in the current time range. The node quantity range in the resource plan has a higher priority.

For example, this parameter is set to **2-4** in this scenario.

- Configuring an Auto Scaling Rule

  To enable **Auto Scaling**, you must configure a scale-out or scale-in rule.

  a. Select **Scale-Out** or **Scale-In**.

  b. Click **Add Rule**. The **Add Rule** page is displayed.

     - You can set a maximum of five rules for scaling out or in a cluster, respectively.
     - The system determines the scale-out and then scale-in based on your configuration sequence. Add important policies first. This prevents repeated scale-out or scale-in operations in case a scale-out or scale-in cannot meet requirements.

  c. Set parameters as required.

Table 9-23 Parameters for adding a rule

Parameter	Description	Example Value
Rule Name	Name of an auto scaling rule	default-expand-2
If	The metric threshold of an auto scaling rule. ▪ For details about auto scaling metrics, see **Table 9-13** in **Automatic Scaling of Task Nodes in an MRS Cluster**. ▪ Comparison factors include greater than, greater than or equal to, less than, and less than or equal to.	YARNAppPending Greater than 75
Last For	The duration (5n) during which a cluster reaches the configured metric threshold for triggering a scale-out or scale-in. The default value of **n** is 1, indicating that the duration is five minutes.	1
Add	Number of Task nodes to be added or removed at a time. The value ranges from 1 to 100.	1
Cooldown Period	Interval between two auto scaling operations. The default cooldown period is 20 minutes. The minimum value is **0**.	20

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

d. Click **OK**.

You can view, edit, or delete the rules you configured in the **Scale-out** or **Scale-in** area on the **Auto Scaling** page.

Step 2 Configure a resource plan.

1. Click **Configure Node Range for Specific Time Range** under **Default Range** or **Add Resource Plan**.
2. Configure **Effective On**, **Time Range**, and **Node Range**.

   Click **Configure Node Range for Specific Time Range** or **Add Resource Plan** to configure multiple resource plans.

   After a resource plan triggers a scale-out or scale-in, auto scaling will not be triggered again within 10 minutes.

Table 9-24 Scaling rule parameters

Parameter	Description	Example Value
Effective On	**Effective On** is set to **Daily** by default. You can also select one or more days from Monday to Sunday.	Monday, Tuesday, and Saturday
Time Range	The start time must be at least 30 minutes earlier than the end time. Time ranges configured for different resource plans cannot overlap.	07:00-13:00
Node Range	– If you set both **Node Range** and **Time Range**, the node range you set will be used during the time range you set, and the default node range will be used beyond the time range you set. – If you do not set **Node Range**, its default value will be used.	5-8

Step 3 Click **OK**.

- If you are configuring a resource plan for an existing cluster, select **I agree to authorize MRS to scale out or in nodes based on the above rule**.
- After the configuration, system automatically scales in or out the Task node group as long as the auto scaling rules are met.
- You can click **Operation Logs** in the navigation pane to view the scaling details of a cluster.

----End

9.5.4.3 Managing MRS Cluster Auto Scaling Policies

You can view, modify, delete, enable, and disable an auto scaling policy on the MRS console after it is created.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Viewing an Auto Scaling Policy

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters**, select a running cluster, and click its name to access its details page.

Step 3 Click the **Auto Scaling** tab.

Step 4 Click **View** on the right of the target auto scaling policy to view it.

----End

Modifying an Auto Scaling Policy

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters**, select a running cluster, and click its name to access its details page.

Step 3 Click the **Auto Scaling** tab.

Step 4 Click **Edit** on the right of the target auto scaling policy.

----End

Deleting an Auto Scaling Policy

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters**, select a running cluster, and click its name to access its details page.

Step 3 Click the **Auto Scaling** tab.

Step 4 Click **Delete** on the right of an existing auto scaling policy.

----End

Enabling or Disabling an Auto Scaling Policy

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters**, select a running cluster, and click its name to access its details page.

Step 3 Click the **Auto Scaling** tab.

Step 4 Toggle **Auto Scaling Policy** on or off to enable or disable an auto scaling policy.

----End

9.6 MRS Cluster Data Backup and Restoration

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.6.1 Backing Up and Restoring MRS Cluster Data

Overview

Manager can back up system and user data by components. The system can back up Manager data, component metadata, and service data.

For MRS 3.*x* and later, data can be backed up to local disks (LocalDir), local HDFS (LocalHDFS), remote HDFS (RemoteHDFS), NAS (NFS/CIFS), SFTP servers (SFTP), and OBS (supported in MRS 3.1.0 and later). For details, see **Backing Up MRS Cluster Component Data**.

Backup and restoration tasks are performed in the following scenarios:

- Routine backup is performed to ensure the data security of the system and components.
- If the system is faulty, the data backup can be used to recover the system.
- If the active cluster is completely faulty, a mirrored cluster identical to the active cluster needs to be created. You can use the backup data to restore the active cluster.

Table 9-25 Metadata (MRS 2.x and earlier versions)

Backup Type	Backup Content
OMS	Database data (excluding alarm data) and configuration data in the cluster management system to be backed up by default
LdapServer	User information, including the username, password, key, password policy, and user group information
DBService	Metadata of the components (Hive) managed by DBService
NameNode	HDFS metadata

Table 9-26 Manager configuration data (MRS 3.x and later)

Backup Type	Backup Content	Backup Directory Type
OMS	Database data (excluding alarm data) and configuration data in the cluster management system by default	• LocalDir • LocalHDFS • RemoteHDFS • NFS • CIFS • SFTP • OBS

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-27 Component metadata or other data (MRS 3.x and later)

Backup Type	Backup Content	Backup Directory Type
DBService	Metadata of the components (including Loader, Hive, Spark, Oozie, CDL, and Hue) managed by DBService.	• LocalDir • LocalHDFS • RemoteHDFS • NFS • CIFS • SFTP • OBS
Flink (Applicable to MRS 3.2.0 and later versions)	Flink metadata.	• LocalDir • LocalHDFS • RemoteHDFS • OBS (available in MRS 3.5.0 and later)
Kafka	Kafka metadata.	• LocalDir • LocalHDFS • RemoteHDFS • NFS • CIFS • OBS
NameNode	HDFS metadata. After multiple NameServices are added, backup and restoration are supported for all of them and the operations are consistent with those of the default hacluster instance.	• LocalDir • RemoteHDFS • NFS • CIFS • SFTP • OBS
Yarn	Information about the Yarn service resource pool.	
HBase	**tableinfo** files and data files of HBase system tables.	
IoTDB	IoTDB metadata.	• LocalDir • NFS • RemoteHDFS • CIFS • SFTP
ClickHouse	ClickHouse metadata.	• LocalDir • RemoteHDFS

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-28 Service data of specific components (MRS 3.x and later)

Backup Type	Backup Content	Backup Directory Type
HBase	Table-level user data.	• RemoteHDFS • NFS • CIFS • SFTP • OBS (available in MRS 3.5.0 and later)
HDFS	Directories or files of user services. Encrypted directories cannot be backed up or restored.	
Hive	Table-level user data.	
IoTDB	IoTDB service data.	RemoteHDFS
ClickHouse	Table-level user data.	RemoteHDFS
Doris	Doris service data. This function is available for MRS 3.5.0 and later.	• RemoteHDFS • OBS

Note that some components in MRS 3.x and later versions do not provide data backup or restoration:

- Kafka supports replicas and allows multiple replicas to be specified when a topic is created.
- For MRS 3.5.0 and later, Kafka, as a message channel, does not store data permanently. By default, only data of the latest seven days is stored. Independent data backup is not supported.
- CDL data is stored in DBService and Kafka. A system administrator can create DBService and Kafka backup tasks to back up data.
- MapReduce and Yarn data is stored in HDFS. Therefore, they rely on the backup and restoration provided by HDFS.
- Backup and restoration of service data in ZooKeeper are performed by their own upper-layer components.

MRS Cluster Data Backup and Restoration Principles

Task

Before backup or restoration, you need to create a backup or restoration task and set task parameters, such as the task name, backup data source, and type of the directory for storing backup files. Then you can execute the tasks to back up or restore data. When Manager is used to restore the data of HDFS, HBase (MRS 3.x and later), Hive, and NameNode, the cluster cannot be accessed.

Each backup task can back up data of different data sources and generate an independent backup file for each data source. All the backup files generated in a backup task form a backup file set, which can be used in restoration tasks. Backup data can be stored on Linux local disks, local cluster HDFS, and standby cluster HDFS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 3.x and later versions, backup tasks support full backup and incremental backup policies. Cloud data backup tasks do not support incremental backup. If the backup directory type is NFS or CIFS, incremental backup is not recommended. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.
- For MRS 2.x and earlier versions, the backup task provides the full backup or incremental backup policies. HDFS and Hive backup tasks support the incremental backup policy, while OMS, LdapServer, DBService, and NameNode backup tasks support only the full backup policy.

Task execution rules:

- If a task is already running, you cannot start it again or launch another task until it finishes.
- The automatic execution interval for periodic tasks must exceed 120 seconds. If the interval is less, task execution is postponed until the subsequent period. Manual tasks can be executed at any interval.
- During automatic execution of a periodic task, the current time should not be more than 120 seconds later than the task's start time. If this condition is violated, task execution is postponed until the subsequent period.
- When a periodic task is locked, it cannot run automatically. You need to manually unlock it before it can start.
- The **LocalBackup** partition on the active management node must have at least 20 GB of free space to start the backup tasks for OMS, LdapServer (MRS 2.x or earlier), DBService, Kafka (MRS 3.x or later), and NameNode.
- When planning backup and restoration tasks, select the data to be backed up or restored strictly based on the service logic, data store structure, and database or table associations.
  - For MRS 2.x and earlier versions, the system automatically creates the default periodic backup task **default** with a 24- hour execution interval. This task performs full backup of OMS, LdapServer, DBService, and NameNode data to the Linux local disk.
  - For MRS 3.x and later versions, the system automatically creates default periodic backup tasks **default-oms** and **default-*cluster ID*** with a 1- hour execution interval. These tasks perform full backup of OMS metadata and cluster metadata, including DBService and NameNode, to local disks.

Snapshot (MRS 3.x and later versions)

The system uses the snapshot technology to quickly back up data. Snapshots include HBase and HDFS snapshots.

- HBase snapshots

  An HBase snapshot is a backup file of HBase tables at a specified time point. This backup file does not replicate service data or affect the RegionServer. The HBase snapshot replicates table metadata, including table descriptor, region info, and HFile reference information. The metadata can be used to restore data before the snapshot creation time.
- HDFS snapshots

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

An HDFS snapshot is a read-only backup of HDFS at a specified time point. The snapshot is used in data backup, misoperation protection, and disaster recovery scenarios.

The snapshot function can be enabled for any HDFS directory to create the related snapshot file. Before creating a snapshot for a directory, the system automatically enables the snapshot function for the directory. Creating a snapshot does not affect any HDFS operation. A maximum of 65,536 snapshots can be created for each HDFS directory.

If a snapshot has been created for an HDFS directory, the directory cannot be deleted or modified until the snapshot is removed. Snapshots cannot be created for the upper-layer directories or subdirectories of the directory.

DistCp (MRS 3.x and later versions)

Distributed copy (DistCp) is a tool used to replicate a large amount of data in HDFS in a cluster or between the HDFSs of different clusters. In a backup or restoration task of HBase, HDFS, or Hive, if you back up the data to HDFS of the standby cluster, the system invokes DistCp to perform the operation. Install the MRS software of the same version for the active and standby clusters and install the cluster.

DistCp uses MapReduce to implement data distribution, troubleshooting, restoration, and report. DistCp specifies different Map jobs for various source files and directories in the specified list. Each Map job copies the data in the partition that corresponds to the specified file in the list.

If you use DistCp to replicate data between HDFSs of two clusters, configure the cross-cluster mutual trust (mutual trust does not need to be configured for clusters managed by the same FusionInsight Manager) and cross-cluster replication for both clusters. When backing up the cluster data to HDFS in another cluster, you need to install the Yarn component. Otherwise, the backup fails.

Local Fast Restoration (MRS 3.x and later versions)

After using DistCp to back up the HBase, HDFS, and Hive data of the local cluster to the HDFS of the standby cluster, the HDFS of the local cluster retains the backup data snapshots. You can create local rapid restoration tasks to restore data by using the snapshot files in the HDFS of the local cluster.

NAS (MRS 3.x and later versions)

Network Attached Storage (NAS) is a dedicated data storage server which includes the storage components and embedded system software. It provides the cross-platform file sharing function. By using NFS (supporting NFSv3 and NFSv4) and CIFS (supporting SMBv2 and SMBv3), you can connect the service plane of MRS to the NAS server to back up data to the NAS or restore data from the NAS.

- Before data is backed up to the NAS, the system automatically mounts the NAS shared address to a local partition on the backup task execution node. After the backup is complete, the system unmounts the NAS shared partition from the backup task execution node.
- To prevent backup and restoration failures, do not access the shared address where the NAS server has been mounted to, for example, **/srv/BigData/LocalBackup/nas**, during data backup and restoration.
- DistCp is used to back up service data to the NAS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Specifications of MRS Cluster Data Backup and Restoration

Table 9-29 Specifications of the backup and restoration feature

Item	Specification
Maximum number of backup or restoration tasks	100
Number of concurrent tasks in a cluster	1
Maximum number of waiting tasks	199
Maximum size (GB) of backup files on a Linux local disk	600

In MRS 3.x and later versions, when service data is stored in ZooKeeper upper- layer components, do not include too many znodes in a single backup or restoration task. Excessive znodes may result in task failure and adversely affect ZooKeeper service performance. To check the number of znodes in a single backup or restoration task, perform the following operations:

- Ensure that the number of znodes included in any backup or restoration task remains below the operating system's maximum file handle threshold. Methods for checking the maximum file handle threshold are as follows:
  a. To check the system-level threshold, run the **cat /proc/sys/fs/file-max** command.
  b. To check the user-level threshold, run the **ulimit -n** command.
- When the number of znodes in a parent directory surpasses the defined threshold, data backup and restoration shall be executed in batches across its sub- directories. To check the number of znodes using ZooKeeper client scripts, perform the following operations:
  a. On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **ZooKeeper**. Click **Instances** and check the management IP address of each ZooKeeper role instance.
  b. Log in to the node where the client is installed, configure environment variables, authenticate the user (skip this operation for clusters with Kerberos authentication disabled), and run the following command:

     zkCli.sh -server *ip:port*

     ip can be any management IP address. The default value of *port* is 2181.
  c. If the following information is displayed, login to the ZooKeeper server is successful:

     ```
     WatchedEvent state:SyncConnected type:None path:null
     [zk: ip:port(CONNECTED) 0]
     ```
  d. Run the **getusage** command to check the number of znodes in the directory to be backed up.

     getusage */hbase/region*

     In the command output, **Node count=xxxxxx** indicates the number of znodes stored in the **region** directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-30 Specifications of the **default** task (MRS 2.x and earlier)

Item	OMS	LdapServer	DBService	NameNode
Backup period	1 hour			
Maximum number of backup copies	2			
Maximum size of a backup file	10 MB	20 MB	100 MB	1.5 GB
Maximum size of disk space used	20 MB	40 MB	200 MB	3 GB
Save path of backup data	*Data save path*/**LocalBackup/** of the active and standby management nodes			

Table 9-31 Specifications of the **default** task (MRS 3.x and later)

Item	OMS	HBase	Kafka	DBService	NameNode
Backup period	1 hour				
Maximum number of backups	168 (7-day historical data)				24 (one-day historical data)
Maximum size of a backup file	10 MB	10 MB	512 MB	100 MB	20 GB
Maximum size of disk space used	1.64 GB	1.64 GB	84 GB	16.41 GB	480 GB
Storage path of backup data	*Data storage path*/**LocalBackup/** of the active and standby management nodes				

- The backup data generated by the default backup task must be periodically transferred and preserved outside the cluster, as mandated by enterprise O&M requirements.
- In MRS 3.x and later, administrators can create DistCp backup tasks to save OMS, DBService, and NameNode data to external clusters.
- The execution time of a cluster data backup task is calculated as the volume of data to be backed up divided by the network bandwidth between the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cluster and the backup device. For practical estimation, the calculated duration should be multiplied by 1.5 to obtain the reference execution time.

- Executing a data backup task affects the maximum I/O performance of the cluster. Therefore, you are advised to execute a backup task during off-peak hours.

9.6.2 Enabling MRS Inter-Cluster Replication

DistCp is used to replicate the data stored in HDFS from a cluster to another cluster. To use DistCp, you must first enable inter-cluster replication on both nodes of the cluster where you want to copy data.

Administrators can modify parameters on Manager to enable inter-cluster replication. They then create a backup task to copy data to the remote HDFS.

Impact on the System

Yarn needs to be restarted to enable the cross-cluster replication function and cannot be accessed during restart.

Prerequisites

- The **hadoop.rpc.protection** parameter of HDFS in the two clusters for data replication must use the same data transmission mode. The default value is **privacy**, indicating encrypted transmission. The value **authentication** indicates that transmission is not encrypted.
- For clusters with Kerberos authentication enabled (security mode), mutual trust between clusters needs to be configured.
- The inbound rules of the two security groups on the peer cluster have been added to the two security groups in each cluster to allow all access requests of all protocols and ports of all ECSs in the security groups.

Enabling MRS Inter-Cluster Replication

Step 1 Log in to the MRS Manager of either cluster.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

- For MRS 2.x and earlier, choose **Services** > **Yarn** > **Service Configuration** and set **Type** to **All**.
- For MRS 3.x and later, choose **Cluster** > **Services** > **Yarn** > **Configurations**, and click **All Configurations**.

Step 2 In the navigation pane on the left, choose **Yarn** > **Distcp** and set the following parameters:

- For MRS2.x and earlier versions, set **dfs.namenode.rpc-address.haclusterX.remotenn1** to the service IP address and RPC port of a NameNode instance in the peer cluster, in **dfs.namenode.rpc-address.haclusterX.remotenn2**, enter the service IP address and RPC port number of the other NameNode instance in the peer cluster. For example, enter **10.1.1.1:25000** and **10.1.1.2:25000**.

  dfs.namenode.rpc-address.haclusterX.remotenn1 and **dfs.namenode.rpc-address.haclusterX.remotenn2** do not distinguish active and standby

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NameNode instances. The default NameNode RPC port is **25000** and cannot be modified on Manager.

- For MRS 3.x and later versions, modify **dfs.namenode.rpc-address**, set **haclusterX.remotenn1** to the service IP address and RPC port of one NameNode instance of the peer cluster, and set **haclusterX.remotenn2** to the service IP address and RPC port number of the other NameNode instance of the peer cluster. Examples of modified parameter values: **10.1.1.1:8020** and **10.1.1.2:8020**.

  haclusterX.remotenn1 and **haclusterX.remotenn2** do not distinguish active and standby NameNodes. The default NameNode RPC port is 8020 and cannot be modified on Manager.

  If data of the current cluster needs to be backed up to the HDFS of multiple clusters, you can configure the corresponding NameNode RPC addresses to **haclusterX1**, **haclusterX2**, **haclusterX3**, and **haclusterX4**.

Step 3 Save the configurations and restart YARN.

- For MRS2.x and earlier versions, click **Save Configuration** and select **Restart the affected services or instances**. Click **OK** to restart the YARN service.

  After the system displays "Operation succeeded", click **Finish**. The YARN service is restarted successfully.
- For MRS 3.x and later versions, click **Save**. In the displayed dialog box, click **OK**. Click **Dashboard**, choose **More** > **Restart Service**, and enter the password of the user to restart the YARN service.

Step 4 Log in to Manager of the other cluster and repeat **Step 1** to **Step 3**.

----End

9.6.3 Creating an MRS Cluster Data Backup Task

You can create backup tasks on Manager. Executing backup tasks backs up related data.

Creating a Data Backup Task (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**. On the page that is displayed, click **Create**.

Step 3 Set **Backup Object** to **OMS** or the cluster whose data you want to back up.

Step 4 Enter a task name in the **Name** text box.

Step 5 Set **Mode** to **Periodic** or **Manual** as required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-32 Backup types

Type	Parameter	Example Value	Description
Periodic backup	Start Time	2024/11/22 10:29:11	Indicates the time when a periodic backup task is started for the first time.
	Period	1 day	Task execution interval. Value options are **Hours** and **Days**.
	Backup Policy	Full backup at the first time and incremental backup subsequently	The following policies can be selected: • **Full backup at the first time and subsequent incremental backup** • **Full backup every time** • **Full backup once every n times**
Manual backup	N/A	-	You need to manually execute the task to back up data.

Step 6 Set required parameters in the **Configuration** area.

- Metadata and service data can be backed up.
- For details about how to back up data of different components, see **MRS Cluster Data Backup and Restoration**.

Step 7 Click **OK** to save the configurations.

Step 8 In the backup task list, you can view the created backup task.

Locate the row that contains the target backup task, choose **More** > **Back Up Now** in the **Operation** column to execute the task immediately.

----End

Creating a Metadata Backup Task (MRS 2.x and Earlier)

Step 1 Create a backup task.

1. Log in to FusionInsight Manager of the MRS cluster.

   For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
2. Choose **System** > **Back Up Data**.
3. Click **Create Backup Task**.

Step 2 Configure a backup policy.

1. Set **Name** to the name of the backup task.
2. Set **Backup Mode** to the type of the backup task. **Periodic** indicates that the backup task is periodically executed. **Manual** indicates that the backup task is manually executed.

   To create a periodic backup task, set the following parameters:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Started**: indicates the time when the task is started for the first time.
- **Period**: indicates the task execution interval. The options include **Hours** and **Days**.
- **Backup Policy**: indicates the volume of data to be backed up in each task execution. The options include **Full backup at the first time and incremental backup later**, **Full backup every time**, and **Full backup once every n times**. If you select **Full backup once every n times**, you need to specify the value of *n*.

Step 3 Select backup sources.

On the **Configuration** page, select the metadata option and set backup parameters. For details, see **Backing Up Manager Data (MRS 2.x and Earlier)**.

Step 4 Click **OK**.

Step 5 In the **Operation** column of the created task in the backup task list, click **Back Up Now** if **Backup Mode** is set to **Periodic** or click **Start** If **Backup Mode** is set to **Manual** to execute the backup task.

----End

9.6.4 Creating an MRS Cluster Data Restoration Task

You can create a restoration task on Manager. After the restoration task is executed, the specified backup data is restored to the cluster.

Creating a Data Restoration Task (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Backup and Restoration** > **Restoration Management**. On the page that is displayed, click **Create**.

Step 3 Configure **Task Name**.

Step 4 Set **Recovery Object** to **OMS** or the cluster whose data you want to restore.

Step 5 Set the required parameters in the **Recovery Configuration** area.

- Metadata and service data can be restored.
- For details about how to restore data of different components, see **MRS Cluster Data Backup and Restoration**.

Step 6 Click **OK** to save the configurations.

Step 7 In the restoration task list, you can view the created restoration tasks.

Locate the row containing the target restoration task, click **Start** in the **Operation** column to execute the restoration task immediately.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Creating a Metadata Restoration Task (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 On MRS Manager, choose **System** > **Recovery Management**.

Step 3 On the page that is displayed, click **Create Restoration Task**.

Step 4 Set **Task Name** to the name of the restoration task.

Step 5 In the **Configuration** area, select the component whose metadata is to be restored and set restoration parameters. For details, see **Restoring Manager Data (MRS 2.x and Earlier)**.

Step 6 Click **OK**.

Step 7 In the restoration task list, you can view the created restoration tasks.

In the restoration task list, locate the row where the created task resides, and click **Start** in the **Operation** column.

----End

9.6.5 Backing Up MRS Cluster Component Data

9.6.5.1 Backing Up Manager Data (MRS 2.x and Earlier)

Scenarios

To ensure the security of metadata either on a routine basis or before and after performing critical metadata operations (such as scale-out, scale-in, patch installation, upgrades, and migration), metadata must be backed up. The backup data can be used to recover the system if an exception occurs or if the operation has not achieved the expected result. This minimizes the adverse impact on services. Metadata includes data of OMS, LdapServer, DBService, and NameNode. MRS Manager data to be backed up includes OMS data and LdapServer data.

By default, metadata backup is supported by the **default** task. This section describes how to back up metadata using MRS Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up Manager data.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **LocalHDFS**: indicates that backup files are stored in the HDFS directory of the current cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- A standby cluster for backing up data has been created, and the network is connected. The inbound rules of the two security groups on the peer cluster have been added to the two security groups in each cluster to allow all access requests of all protocols and ports of all ECSs in the security groups.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.

Backing Up Manager Data

Step 1 Create a backup task.

1. Log in to FusionInsight Manager of the MRS cluster.

   For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.
2. Choose **System** > **Back Up Data**.
3. Click **Create Backup Task**.

Step 2 Configure a backup policy.

1. Set **Task Name** to the name of the backup task.
2. Set **Backup Mode** to the type of the backup task. **Periodic** indicates that the backup task is periodically executed. **Manual** indicates that the backup task is manually executed.

   To create a periodic backup task, set the following parameters:

   - **Started**: indicates the time when the task is started for the first time.
   - **Period**: indicates the task execution interval. The options include **By hour** and **By day**.
   - **Backup Policy**: indicates the volume of data to be backed up in each task execution. The options include **Full backup at the first time and incremental backup later**, **Full backup every time**, and **Full backup once every n times**. If you select **Full backup once every n times**, you need to specify the value of *n*.

Step 3 Select backup sources.

In the **Configuration** area, select **OMS** and **LdapServer** under **Metadata**.

Step 4 Set backup parameters.

1. Set **Path Type** of **OMS** and **LdapServer** to a backup directory type.

Table 9-33 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
LocalHDFS	Target Path	HDFS directory for storing the backup files. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	New Instance Name	NameService name of the backup directory. The default value is **hacluster**.

2. Click **OK**.

Step 5 Execute the backup task.

In the **Operation** column of the created task in the backup task list, click **Back Up Now** if **Backup Mode** is set to **Periodic** or click **Start** If **Backup Mode** is set to **Manual** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup Manager data may be restored if data is accidentally modified, deleted, or requires retrieval; if major metadata component operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring Manager Data (MRS 2.x and Earlier)**.

9.6.5.2 Backing Up Manager Data (MRS 3.x and Later Versions)

Scenarios

To ensure data security of FusionInsight Manager routinely or before and after a critical operation (such as capacity expansion and reduction) on FusionInsight Manager, you need to back up FusionInsight Manager data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

Manager data backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up Manager data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **LocalHDFS**: indicates that backup files are stored in the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

MRS 3.1.0 and later versions support the backup of Manager data to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

Backing Up Manager Data

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Set **Backup Object** to **OMS**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-34 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup every time**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **OMS**.

Step 8 Select a backup directory type for **Path Type** of **OMS**.

Table 9-35 Backup directory types

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
LocalHDFS	Target Path	HDFS directory for storing the backup files. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Cluster for Backup	Cluster name of the backup directory.
	Target NameService Name	NameService name of the backup directory. The default value is **hacluster**.
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster, or to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Source Cluster	Cluster containing the YARN queue used by the data you want to back up.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
OBS	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup Manager data may be restored if data is accidentally modified, deleted, or requires retrieval; if major data adjustments on FusionInsight Manager cause exceptions in system data or fail to achieve the expected result; or if all modules fail and become unavailable. For detailed operations, refer to **Restoring Manager Data (MRS 3.x and Later Versions)**.

9.6.5.3 Backing Up CDL Service Data

Scenarios

To ensure CDL service data security routinely or before a major operation on CDL (such as upgrade or migration), you need to back up CDL data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

CDL data is stored in DBService and Kafka. You can create backup tasks for DBService and Kafka on FusionInsight Manager. The system supports both automatic and manual backup modes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

Backing Up CDL Service Data

Step 1 Log in to MRS Manager.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-36 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Backup Policy	**Full backup every time**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 Set **Configuration** to **DBService** and **Kafka**.

Step 8 Select a backup directory type for **Path Type** of **DBService**. For details about how to set the parameters, see **Step 8**.

Step 9 Select a backup directory type for **Path Type** of **Kafka**. For details about how to set the parameters, see **Step 8**.

Step 10 Click **OK**.

Step 11 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup CDL service data may be restored if data is accidentally modified, deleted, or requires retrieval; if major CDL operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring CDL Service Data**.

9.6.5.4 Backing Up ClickHouse Metadata

Scenarios

To ensure ClickHouse metadata security or before a major operation (such as upgrade or migration), you need to back up ClickHouse metadata. The backup

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

ClickHouse metadata backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up ClickHouse metadata.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

- MRS 3.1.0 or later supports this function.
- MRS 3.3.0 and later versions support the backup of ClickHouse metadata to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
  - In active/standby mode, ensure that the value of **HADOOP_RPC_PROTECTION** of ClickHouse is the same as that of **hadoop.rpc.protection** of HDFS.
  - When data is remotely backed up to HDFS, HDFS encrypted directories are not supported.
  - If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.

Backing Up ClickHouse Metadata

Step 1 Log in to MRS Manager.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task. **Periodic** indicates that the backup task is periodically executed. **Manual** indicates that the backup task is manually executed.

To create a periodic backup task, set the following parameters:

- **Started**: indicates the time when the task is started for the first time.
- **Period**: indicates the task execution interval. The options include **Hours** and **Days**.
- **Backup Policy**: Only **Full backup every time** is supported.

Step 7 In **Configuration**, select **ClickHouse** under **Metadata and other data**.

Step 8 Select a backup directory type for **Path Type** of **ClickHouse**.

If **Path Type** is **RemoteHDFS** and the version is MRS 3.1.0 or MRS 3.1.2, configure the environment by referring to **Configuring the Environment When the ClickHouse Backup Task's Path is RemoteHDFS**.

Table 9-37 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Destination Active NameNode IP Address	Service plane IP address of the active NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the active NameNode. This parameter is available only for MRS 3.2.0 and later clusters.
	Destination Standby NameNode IP Address	Service plane IP address of the standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the standby NameNode. This parameter is available only for MRS 3.2.0 and later clusters.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the NameNode. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
	Destination NameNode RPC Port	Value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the destination cluster. This parameter is available only for MRS 3.2.0 and later clusters.
	Target Path	HDFS directory in the standby cluster for storing backup data. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
OBS	Target Path	OBS directory for storing backup data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Data source_Task execution time*.**tar.gz**.

----End

Configuring the Environment When the ClickHouse Backup Task's Path is RemoteHDFS

This topic is available for MRS 3.1.0 only.

Step 1 Log in to MRS Manager of the standby cluster.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > **HDFS**. Click **More** and select **Download Client**. Set **Select Client Type** to **Configuration Files Only**, select **x86_64** for x86 or **aarch64** for Arm based on the type of the node, and click **OK**.

Step 3 After the client file package is generated, download the client to the local PC as prompted and decompress the package.

For example, if the client file package is **FusionInsight_Cluster_1_HDFS_Client.tar**, decompress it to obtain **FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles.tar**. Then, decompress this file to the **D:\FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles** directory on the local PC. The directory name cannot contain spaces.

Step 4 Go to the **FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles** client directory and obtain the **hosts** file.

Step 5 Go to **FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles\HDFS\config** to obtain the **core-site.xml** and **hdfs-site.xml** files.

Step 6 Log in to MRS Manager of the source cluster.

Step 7 Choose **Cluster** > **Services** > **ClickHouse**, click **Configurations** and then **All Configurations**, and select **Backup** under **ClickHouse(Service)**.

For **remote_connect_core_config_file**, click **Upload File** and select the **core-site.xml** file prepared in **Step 5**.

For **remote_connect_hdfs_config_file**, click **Upload File** and select the **hdfs-site.xml** file prepared in **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click **Save**, confirm the information, and click **OK** to save the configuration. After saving the configurations, click **Finish**.

Step 9 Choose **Cluster** > **Services** > **ClickHouse**, click **Instances**, and view the instance IP address of **ClickHouseServer**.

Step 10 Log in to the nodes of the ClickHouseServer instances as user **root** and check whether the **/etc/hosts** file contains the host information in **Step 4**. If not, add the host information in **Step 4** to the **/etc/hosts** file.

----End

Helpful Links

Backup ClickHouse metadata may be restored if data is accidentally modified, deleted, or requires retrieval; if major ClickHouse operations (such as upgrades or migration) cause exceptions in system data or fail to achieve the expected result; if all ClickHouse modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring ClickHouse Metadata**.

9.6.5.5 Backing Up ClickHouse Service Data

Scenarios

To ensure ClickHouse service data security routinely or before a major operation on ClickHouse (such as upgrade or migration), you need to back up ClickHouse service data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

ClickHouse service data backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up ClickHouse service data.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

- MRS 3.1.0 or later supports this function.
- MRS 3.3.0-LTS.1 and later patch versions support the backup of ClickHouse service data to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager,

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.

- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.
- In active/standby mode, ensure that the value of **HADOOP_RPC_PROTECTION** of ClickHouse is the same as that of **hadoop.rpc.protection** of HDFS.
- When data is remotely backed up to HDFS, HDFS encrypted directories are not supported.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

- You have planned the backup type, period, object, and directory based on service requirements.

Backing Up ClickHouse Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-38 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Backup Policy	**Full backup at the first time and incremental backup subsequently**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported.

Step 7 In **Configuration**, select **ClickHouse** under **Service Data**.

Step 8 Select a backup directory type for **Path Type** of **ClickHouse**.

If **Path Type** is **RemoteHDFS** and the version is MRS 3.1.0 or MRS 3.1.2, configure the environment by referring to **Configuring the Environment When the ClickHouse Backup Task's Path is RemoteHDFS**.

Table 9-39 Path of backup data

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Destination Active NameNode IP Address	Service plane IP address of the active NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the active NameNode. This parameter is available only for MRS 3.2.0 and later clusters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Destination Standby NameNode IP Address	Service plane IP address of the standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the standby NameNode. This parameter is available only for MRS 3.2.0 and later clusters.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the NameNode. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
	Destination NameNode RPC Port	Value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the destination cluster. This parameter is available only for MRS 3.2.0 and later clusters.
	Target Path	HDFS directory in the standby cluster for storing backup data. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
OBS	Target Path	OBS directory for storing backup data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Set **Maximum Number of Recovery Points** to the number of snapshots that can be retained in the cluster.

Step 10 Set **Backup Content** to one or multiple ClickHouse tables to be backed up.

You can select backup data using either of the following methods:

- Adding a backup data file

  Click the name of a database in the navigation pane to show all the tables in the database, and select specified tables.

  MRS 3.2.0 or later:

  a. Click **Add**.

  b. Select the table to be backed up under **File Directory**, and click **Add** to add the table to **Backup Content**.

  c. Click **OK**.

- MRS 3.2.0 or later: Regular expression filtering

  a. Click **Query Regular Expression**.

  b. Enter the logical cluster and database to which the ClickHouse table belongs in the first text box as prompted. The logical cluster and database must match the existing logical cluster and database, for example, **/default_cluster/database**.

  c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to filter all tables that contain the keyword **test** in the database, enter **test.***.

  d. Click **Refresh** to view the displayed tables in **Directory Name**.

  e. Click **Synchronize** to save the result.

  - **When entering regular expressions, click + or − to add or delete an expression.**
  - **If the selected table or directory is incorrect, click Clear Selected Node to deselect it.**

- Versions earlier than MRS 3.2.0: Regular expression filtering

  a. Click **Query Regular Expression**.

  b. In the first text box, enter the name of the database where the ClickHouse tables reside. Ensure that the database is an existing database. for example, **default**.

  c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to get all tables in the database, enter **([\s\S]*?)**. To get tables whose names consist of letters and digits, for example, **tb***1*, enter **tb\d***.

  d. Click **Refresh** to view the displayed tables in **Directory Name**.

  e. Click **Synchronize** to save the result.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **When entering regular expressions, click + or − to add or delete an expression.**
- **If the selected table or directory is incorrect, click Clear Selected Node to deselect it.**

Step 11 Click **Verify** to check whether the backup task is configured correctly.

The possible causes of the verification failure are as follows:

- The target NameNode IP address is incorrect.
- The directory or table to be backed up does not exist.
- The name of the NameService is incorrect.

Step 12 Click **OK**.

Step 13 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Data source_Task creation time*, and the subdirectory is used to save latest data source backup files.

----End

Configuring the Environment When the ClickHouse Backup Task's Path is RemoteHDFS

This topic is available for MRS 3.1.0 only.

Step 1 Log in to FusionInsight Manager of the standby cluster.

Step 2 Choose **Cluster** > **Services** > **HDFS**. Click **More** and select **Download Client**. Set **Select Client Type** to **Configuration Files Only**, select **x86_64** for x86 or **aarch64** for Arm based on the type of the node, and click **OK**.

Step 3 After the client file package is generated, download the client to the local PC as prompted and decompress the package.

For example, if the client file package is **FusionInsight_Cluster_1_HDFS_Client.tar**, decompress it to obtain **FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles.tar**. Then, decompress this file to the **D:\FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles** directory on the local PC. The directory name cannot contain spaces.

Step 4 Go to the **FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles** client directory and obtain the **hosts** file.

Step 5 Go to **FusionInsight_Cluster_1_HDFS_ClientConfig_ConfigFiles\HDFS\config** to obtain the **core-site.xml** and **hdfs-site.xml** files.

Step 6 Log in to FusionInsight Manager of the source cluster.

Step 7 Choose **Cluster** > **Services** > **ClickHouse**, choose **Configurations** > **All Configurations**, and select **backup** under **ClickHouse(Service)**.

For **remote_connect_core_config_file**, click **Upload File** and select the **core-site.xml** file prepared in **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For **remote_connect_hdfs_config_file**, click **Upload File** and select the **hdfs-site.xml** file prepared in **Step 5**.

Step 8 Click **Save**, confirm the information, and click **OK** to save the configuration. After saving the configurations, click **Finish**.

Step 9 Choose **Cluster** > **Services** > **ClickHouse**, click **Instances**, and view the instance IP address of **ClickHouseServer**.

Step 10 Log in to the host nodes of the ClickHouseServer instances as user **root** and check whether the **/etc/hosts** file contains the host information in **Step 4**. If not, add the host information in **Step 4** to the **/etc/hosts** file.

----End

Helpful Links

Backup ClickHouse service data may be restored if data is accidentally modified, deleted, or requires retrieval; if major ClickHouse operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring ClickHouse Service Data**.

9.6.5.6 Backing Up DBService Data

Scenarios

To ensure DBService service data security routinely or before a major operation on DBService (such as upgrade or migration), you need to back up DBService data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

DBService metadata backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up DBService metadata.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **LocalHDFS**: indicates that backup files are stored in the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.
- **OBS**: indicates that backup files are stored in OBS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Notes and Constraints

MRS 3.1.0 and later versions support the backup of DBService metadata to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

Backing Up DBService Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-40 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup every time**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **DBService**.

Step 8 Select a backup directory type for **Path Type** of **DBService**.

Table 9-41 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
LocalHDFS	Target Path	HDFS directory for storing the backup files. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Target NameService Name	NameService name of the backup directory. The default value is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
OBS	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup DBService metadata may be restored if data is accidentally modified, deleted, or requires retrieval; if major DBService operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring DBService Metadata**.

9.6.5.7 Backing Up Doris Data

Scenarios

To ensure Doris service data security, you need to back up Doris service data especially before you perform a major operation on Doris (such as upgrade or migration). The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

Doris service data backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up Doris service data.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

This section applies only to MRS 3.5.0 or later clusters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - At least one DBroker instance of the Doris service has been deployed in the active cluster.
  - If the active and standby clusters are deployed in security mode and they are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active and standby clusters are deployed in normal mode, mutual trust is not required.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
  - The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.
  - The value of **hadoop.rpc.protection** of Doris must be the same as that of **hadoop.rpc.protection** of HDFS in both active and standby clusters.
- You have planned the backup type, period, object, and directory based on service requirements.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

Backing Up Doris Data

Step 1 Log in to FusionInsight Manager of the MRS cluster.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

- **Periodic**: indicates that the backup task is periodically executed. If you select this mode, you need to set other parameters by referring to **Table 9-42**.
- **Manual**: indicates that the backup task is manually executed.

Table 9-42 Periodic backup parameters

Parameter	Description
Started	Time when the task should start for the first time.
Period	Task execution interval. Supported values are **Hours** and **Days**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Backup Policy	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Currently, Doris supports only **full backup each time**. Incremental backup is not supported.

Step 7 In **Configuration**, select **Doris** under **Metadata and other data**.

Step 8 Select a backup directory type for **Path Type** of **Doris**.

Table 9-43 Backup directory types

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active NameNode in the standby cluster.
	Destination NameNode RPC Port	Value of **dfs.namenode.rpc.port** in the HDFS configuration of the destination cluster.
	DBroker IP	Service plane IP address of the DBroker role in the cluster. The DBroker is used to transmit data during backup.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
OBS	Target Path	OBS directory for storing backup data.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Set **Maximum Number of Recovery Points** to the number of snapshots that can be retained in the cluster.

Step 10 Set **Backup Content** to one or multiple Doris tables to be backed up.

You can select backup data using either of the following methods:

- Adding a backup data file

  Click the name of a database in the navigation pane to show all the tables in the database, and select specified tables.

- Selecting using regular expressions

  a. Click **Query Regular Expression**.

  b. Enter the database where the Doris tables are located in the first text box as prompted. The database must be the same as the existing database, for example, **/example_db**.

  c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to search for all tables that contain the keyword **test** in the database, enter **test.***.

  d. Click **Refresh** to view the displayed tables in **Directory Name**.

  e. Click **Synchronize** to save the result.

  – When entering regular expressions, click + or − to add or delete an expression.

  – If the selected table or directory is incorrect, click **Clear Selected Node** to deselect it.

Step 11 Click **Verify** to check whether the backup task is configured correctly.

The possible causes of the verification failure are as follows:

- The target NameNode IP address is incorrect.
- The name of the NameService is incorrect.
- The table to be backed up does not exist.
- The format of the table to be backed up is incorrect.
- The tables to be backed up must come from the same database.

Step 12 Click **OK**.

Step 13 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Data source_Task creation time*, and the subdirectory is used to save latest data source backup files. All the backup file sets are stored in the related snapshot directories.

----End

Helpful Links

Backup Doris service data may be restored if data is accidentally modified, deleted, or requires retrieval; if major Doris operations (such as upgrades or

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring Doris Service Data**.

9.6.5.8 Backing Up Flink Metadata

Scenarios

To ensure Flink metadata security or before a major operation on Flink (such as upgrade or migration), you need to back up Flink metadata. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

Flink metadata backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up Flink metadata.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **LocalHDFS**: indicates that backup files are stored in the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

MRS 3.5.0 and later versions support the backup of Flink metadata to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The HDFS and Yarn services have been installed if data needs to be backed up to HDFS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS. Only MRS 3.5.0 or later supports data backup to OBS.

Backing Up Flink Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task. **Periodic** indicates that the backup task is periodically executed. **Manual** indicates that the backup task is manually executed.

Table 9-44 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup every time**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **Flink** under **Metadata and other data**.

Step 8 Select a backup directory type for **Path Type** of **Flink**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-45 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
LocalHDFS	Target Path	HDFS directory for storing the backup files. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Target NameService Name	NameService name of the backup directory. The default value is **hacluster**.
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
OBS (available in MRS 3.5.0 and later)	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup Flink metadata may be restored if data is accidentally modified, deleted, or requires retrieval; if major Flink operations (such as upgrades or migration) cause exceptions in system data or fail to achieve the expected result; if all Flink modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring Flink Metadata**.

9.6.5.9 Backing Up HBase Metadata

Scenarios

To ensure HBase metadata security (including tableinfo files and HFiles) or before a major operation on HBase system tables (such as upgrade or migration), you need to back up HBase metadata to prevent HBase service unavailability caused by HBase system table directory or file damages. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

HBase metadata backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up HBase metadata.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

MRS 3.1.0 and later versions support the backup of HBase metadata to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- The **fs.defaultFS** parameter settings of HBase are the same as those of Yarn and HDFS.
- If HBase data is stored in the local HDFS, HBase metadata can be backed up to OBS. If HBase data is stored in OBS, data backup is not supported.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

Backing Up HBase Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-46 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup every time**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **HBase** under **Metadata and other data**.

Step 8 Select a backup directory type for **Path Type** of **HBase**.

Table 9-47 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
OBS (available in MRS 3.1.0 and later)	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup HBase metadata may be restored if data is accidentally modified, deleted, or requires retrieval; if major HBase operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring HBase Metadata**.

9.6.5.10 Backing Up HBase Service Data

Scenarios

For HBase service data security, you need to back up HBase service data before a major operation on HBase (such as upgrade or migration). The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

HBase service data backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up HBase service data.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.

The following situations may occur during the HBase service data backup:

- When a user creates an HBase table, **KEEP_DELETED_CELLS** is set to **false** by default. When the user backs up this HBase table, deleted data will be backed up and junk data may exist after data restoration. This parameter can be set

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

to **true** manually when an HBase table is created based on service requirements.

- When a user manually specifies the timestamp when writing data into an HBase table and the specified time is earlier than the last backup time of the HBase table, new data may not be backed up in incremental backup tasks.
- The HBase backup function cannot back up the access control lists (ACLs) for reading, writing, executing, creating, and managing HBase global or namespaces. After HBase data is restored, you need to reset the role permissions on FusionInsight Manager.
- If the backup data of the standby cluster is lost in an existing HBase backup task, the next incremental backup will fail, and you need to create an HBase backup task again. However, the next full backup task will be normal.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
  - The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.
- Backup policies, including the backup task type, period, backup object, backup directory, and Yarn queue required by the backup task are planned based on service requirements.
- On the HDFS client, you have executed the **hdfs lsSnapshottableDir** command as user **hdfs** to check the list of directories for which HDFS snapshots have been created in the current cluster and ensured that the HDFS parent directory or subdirectory where data files to be backed up are stored does not have HDFS snapshots. Otherwise, the backup task cannot be created.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- The **fs.defaultFS** parameter settings of HBase are the same as those of Yarn and HDFS.

Backing Up HBase Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-48 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup at the first time and incremental backup subsequently**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, choose **HBase** > **HBase** under **Service data**.

Step 8 Select a backup directory type for **Path Type** of **HBase**.

Table 9-49 Path of backup data

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

Step 9 Set **Maximum Number of Recovery Points** to the number of snapshots that can be retained in the cluster.

Step 10 Set **Backup Content** to one or multiple HBase tables to be backed up.

You can select backup data using either of the following methods:

- Adding a backup data file

  Click the name of a database in the navigation pane to show all the tables in the database, and select specified tables.

  MRS 3.2.0 or later:

  a. Click **Add**.

  b. Select the table to be backed up under **File Directory**, and click **Add** to add the table to **Backup Content**.

  c. Click **OK**.

- Selecting using regular expressions

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

a. Click **Query Regular Expression**.
b. Enter the namespace where the HBase tables are located in the first text box as prompted. The namespace must be the same as the existing namespace, for example, **default**.
c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to get all tables in the namespace, enter **([\s\S]*?)**. To get tables whose names consist of letters and digits, for example, **tb***1*, enter **tb\d***.
d. Click **Refresh** to view the displayed tables in **Directory Name**.
e. Click **Synchronize** to save the result.

- **When entering regular expressions, click + or − to add or delete an expression.**
- **If the selected table or directory is incorrect, click Clear Selected Node to deselect it.**

Step 11 Click **Verify** to check whether the backup task is configured correctly.

The possible causes of the verification failure are as follows:

- The target NameNode IP address is incorrect.
- The queue name is incorrect.
- The parent directory or subdirectory of the HDFS directory where HBase table data files to be backed up are stored has HDFS snapshots.
- The directory or table to be backed up does not exist.

Step 12 Click **OK**.

Step 13 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *xxx/Backup task name_Data source_Task creation time*, and the subdirectory is used to save latest data source backup files. All the backup file sets are stored in the related snapshot directories.

----End

Helpful Links

Backup HBase service data may be restored if data is accidentally modified, deleted, or requires retrieval; if major HBase operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring HBase Service Data**.

9.6.5.11 Backing Up HDFS NameNode Data

Scenarios

To ensure NameNode service data security routinely or before a major operation on NameNode (such as upgrade or migration), you need to back up NameNode

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

NameNode metadata backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up NameNode metadata.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

MRS 3.1.0 and later versions support the backup of NameNode metadata to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Backing Up HDFS NameNode Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-50 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup every time**	Only **Full backup every time** is supported. • Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. • If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **NameNode**.

Step 8 Select a backup directory type for **Path Type** of **NameNode**.

Table 9-51 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the NameNode in the standby cluster.
	Target Path	Path for storing backup files.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
OBS	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup HDFS NameNode metadata may be restored if data is accidentally modified, deleted, or requires retrieval; if major NameNode operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring HDFS NameNode Metadata**.

9.6.5.12 Backing Up HDFS Service Data

Scenarios

To ensure HDFS service data security routinely or before a major operation on HDFS (such as upgrade or migration), you need to back up HDFS service data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

HDFS service data backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up HDFS service data.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

- Encrypted directories cannot be backed up or restored.
- MRS 3.5.0 and later versions support the backup of HDFS service data to OBS.
- When data is backed up to storage media (such as OBS, SFTP, and CIFS) other than HDFS, the storage media do not support attributes such as the storage duration, owner group, and quota. Therefore, the preceding attributes cannot be restored after the restoration.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
  - The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.
- Backup policies, including the backup task type, period, backup object, backup directory, and Yarn queue required by the backup task are planned based on service requirements.
- On the HDFS client, you have executed the **hdfs lsSnapshottableDir** command as user **hdfs** to check the list of directories for which HDFS snapshots have been created in the current cluster and ensured that the HDFS parent directory or subdirectory where data files to be backed up are stored does not have HDFS snapshots. Otherwise, the backup task cannot be created.
- To back up data to NAS, the following conditions must be met. Otherwise, the backup task may fail.
  - The NAS server has been deployed. The IP address of the NAS server is accessible to all nodes in the cluster, and NAS disks can be mounted on each NodeManager node.
  - The number of concurrent NAS tasks must be greater than or equal to the maximum number of maps configured for the backup task. For details about the parameters, see **Step 8**.
  - The NAS bandwidth must be greater than or equal to the product of the maximum number of maps and the maximum bandwidth of a single map (MB/s). For details about the parameters, see **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS. Only MRS 3.5.0 or later supports data backup to OBS.

Backing Up HDFS Service Data

Step 1 Log in to FusionInsight Manager of the MRS cluster.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-52 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup at the first time and incremental backup subsequently**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **HDFS**.

Step 8 Select a backup directory type for **Path Type** of **HDFS**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-53 Path of backup data

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Server IP Address	IP address of the NAS server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
OBS (available in MRS 3.5.0 and later)	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

Step 9 Set **Maximum Number of Recovery Points** to the number of snapshots that can be retained in the cluster.

Step 10 Set **Backup Content** to one or multiple HDFS directories to be backed up based on service requirements.

You can select backup data using either of the following methods:

- Adding a backup data file

  Click the name of a database in the navigation pane to show all the tables in the database, and select specified tables.

  MRS 3.2.0 or later:

  a. Click **Add**.

  b. Select the table to be backed up under **File Directory**, and click **Add** to add the table to **Backup Content**.

  c. Click **OK**.

- Selecting using regular expressions

  a. Click **Query Regular Expression**.

  b. Enter the parent directory full path of the directory in the first text box as prompted. The directory must be the same as the existing directory, for example, **/tmp**.

  c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to get all files or subdirectories

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

in the parent directory, enter **([\s\S]*?)**. To get files whose names consist of letters and digits, for example, **file*1***, enter **file\d***.

d. Click **Refresh** to view the displayed directories in **Directory Name**.

e. Click **Synchronize** to save the result.

- **When entering regular expressions, click + or − to add or delete an expression.**
- **If the selected table or directory is incorrect, click Clear Selected Node to deselect it.**
- The backup directory must not contain files that have been written for a long time. Otherwise, the backup task will fail. Therefore, you are not advised to perform operations on top-level directories such as **/user**, **/tmp**, and **/mr-history**.

Step 11 Click **Verify** to check whether the backup task is configured correctly.

The possible causes of the verification failure are as follows:

- The target NameNode IP address is incorrect.
- The queue name is incorrect.
- The parent directory or subdirectory of the HDFS directory where data files to be backed up are stored has HDFS snapshots.
- The directory or table to be backed up does not exist.
- The name of the NameService is incorrect.

Step 12 Click **OK**.

Step 13 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Data source_Task creation time*, and the subdirectory is used to save latest data source backup files. All the backup file sets are stored in the related snapshot directories.

----End

Helpful Links

Backup HDFS service data may be restored if data is accidentally modified, deleted, or requires retrieval; if major HDFS operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring HDFS Service Data**.

9.6.5.13 Backing Up Hive Service Data

Scenarios

To ensure Hive service data security routinely or before a major operation on Hive (such as upgrade or migration), you need to back up Hive service data. The

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

Hive service data backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

- Hive backup and restoration cannot identify the service and structure relationships of objects such as Hive tables, indexes, and views. When executing backup and restoration tasks, you need to manage unified restoration points based on service scenarios to ensure proper service running.
- Hive backup and restoration do not support Hive on RDB data tables. You need to back up and restore original data tables in external databases independently.
- If the backup data of the standby cluster is lost in an existing Hive backup task that contains Hive on HBase tables, the next incremental backup will fail, and you need to create a Hive backup task again. However, the next full backup task will be normal.
- After the backup function of FusionInsight Manager is used to back up the HDFS directories at the Hive table level, the Hive tables cannot be deleted and recreated.

MRS clusters support multiple data path types for backing up Hive service data.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

MRS 3.5.0 and later versions support the backup of Hive service data to OBS.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
  - The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Backup policies, including the backup task type, period, backup object, backup directory, and Yarn queue required by the backup task are planned based on service requirements.
- On the HDFS client, you have executed the **hdfs lsSnapshottableDir** command as user **hdfs** to check the list of directories for which HDFS snapshots have been created in the current cluster and ensured that the HDFS parent directory or subdirectory where data files to be backed up are stored does not have HDFS snapshots. Otherwise, the backup task cannot be created.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS. Only MRS 3.5.0 or later supports data backup to OBS.

Backing Up Hive Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-54 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Backup Policy	**Full backup at the first time and incremental backup subsequently**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, choose **Hive** > **Hive**.

Step 8 Select a backup directory type for **Path Type** of **Hive**.

Table 9-55 Path of backup data

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Server Shared Path	Backup path on the SFTP server.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
	NameService Name	NameService name of the backup directory. The default value is **hacluster**.
OBS	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Set **Maximum Number of Recovery Points** to the number of snapshots that can be retained in the cluster.

Step 10 Set **Backup Content** to one or multiple Hive tables to be backed up.

You can select backup data using either of the following methods:

- Adding a backup data file

  Click the name of a database in the navigation pane to show all the tables in the database, and select specified tables.

  MRS 3.2.0 or later:

  a. Click **Add**.
  b. Select the table to be backed up under **File Directory**, and click **Add** to add the table to **Backup Content**.
  c. Click **OK**.

- Selecting using regular expressions

  a. Click **Query Regular Expression**.
  b. In the first text box, enter the name of the database where the Hive tables reside. Ensure that the database is an existing database. For example, enter **default**.
  c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to get all tables in the database, enter **([\s\S]*?)**. To get tables whose names consist of letters and digits, for example, **tb1**, enter **tb\d***.
  d. Click **Refresh** to view the displayed tables in **Directory Name**.
  e. Click **Synchronize** to save the result.

  – **When entering regular expressions, click + or − to add or delete an expression.**
  – **If the selected table or directory is incorrect, click Clear Selected Node to deselect it.**

Step 11 Click **Verify** to check whether the backup task is configured correctly.

The possible causes of the verification failure are as follows:

- The target NameNode IP address is incorrect.
- The queue name is incorrect.
- The parent directory or subdirectory of the HDFS directory where data files to be backed up are stored has HDFS snapshots.
- The directory or table to be backed up does not exist.
- The name of the NameService is incorrect.

Step 12 Click **OK**.

Step 13 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

is *Backup task name_Data source_Task creation time*, and the subdirectory is used to save latest data source backup files. All the backup file sets are stored in the related snapshot directories.

----End

Helpful Links

Backup Hive service data may be restored if data is accidentally modified, deleted, or requires retrieval; if major Hive operations (such as upgrades or significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring Hive Service Data**.

9.6.5.14 Backing Up IoTDB Metadata

Scenarios

To ensure IoTDB metadata security and prevent the IoTDB service from being unavailable due to IoTDB metadata file damages, you need to back up IoTDB metadata. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

You can create a backup task on FusionInsight Manager to back up IoTDB metadata. Both automatic and manual backup tasks are supported.

Prerequisites

- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- Backup policies, including the backup task type, period, backup object, backup directory, and Yarn queue required by the backup task are planned based on service requirements.
- The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.

Backing Up IoTDB Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-56 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup every time**	Indicates a periodic backup policy. ● Full backup at the first time and incremental backup subsequently ● Full backup every time ● Full backup once every *n* times Incremental backup is not supported for component metadata. Only **Full backup every time** is supported.

Step 7 In **Configuration**, select **IoTDB** under **Metadata and other data**.

Step 8 Select a backup directory type for **Path Type** of **IoTDB**.

The following backup directory types are supported:

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files. By default, the backup files are stored in *Data storage path*/**LocalBackup/**.

  If you select this option, you need to set the maximum number of replicas to specify the number of backup file sets that can be retained in the backup directory.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.

  If you select this option, set the following parameters:

  – **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.

  – **Server IP Address**: indicates the IP address of the NAS server.

  – **Server Shared Path**: indicates the configured shared directory of the NAS server. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.

  – **Maximum Number of Backup Copies**: indicates the number of backup file sets that can be retained in the backup directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.

  If you select this option, set the following parameters:

  - **Destination NameService Name**: indicates the NameService name of the standby cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
  - **Destination Active NameNode IP Address**: indicates the service plane IP address of the active NameNode in the standby cluster.
  - **Destination Standby NameNode IP Address**: indicates the service plane IP address of the standby NameNode in the standby cluster.
  - **Destination NameNode RPC Port**: indicates the value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the standby cluster.
  - **Target Path**: indicates the HDFS directory for storing standby cluster backup data. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.
  - **Maximum Number of Backup Copies**: indicates the number of backup file sets that can be retained in the backup directory.

- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.

  If you select this option, set the following parameters:

  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
  - **Server IP Address**: indicates the IP address of the NAS server.
  - **Port**: indicates the port number used to connect to the NAS server over the CIFS protocol. The default value is **445**.
  - **Username**: indicates the username set when the CIFS protocol is configured.
  - **Password**: indicates the password set when the CIFS protocol is configured.
  - **Server Shared Path**: indicates the configured shared directory of the NAS server. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
  - **Maximum Number of Backup Copies**: indicates the number of backup file sets that can be retained in the backup directory.

- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.

  If you select this option, set the following parameters:

  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Server IP Address**: indicates the IP address of the server where the backup data is stored.
- **Port**: indicates the port number used to connect to the backup server over the SFTP protocol. The default value is **22**.
- **Username**: indicates the username for connecting to the server using the SFTP protocol.
- **Password**: indicates the password for connecting to the server using the SFTP protocol.
- **Server Shared Path**: indicates the backup path on the SFTP server.
- **Maximum Number of Backup Copies**: indicates the number of backup file sets that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

9.6.5.15 Backing Up IoTDB Service Data

Scenarios

To ensure IoTDB service data security routinely or before a major operation on IoTDB (such as upgrade or migration), you need to back up IoTDB service data. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

You can create a backup task on FusionInsight Manager to back up IoTDB service data. Both automatic and manual backup tasks are supported.

Prerequisites

To back up IoTDB service data to a remote HDFS, you need to meet the following conditions:

- A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
- If the cluster where IoTDB is deployed is in normal mode, service data cannot be backed up to the HDFS of a security-mode cluster.
- If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The HDFS in the standby cluster has sufficient space. You are advised to save backup files in a custom directory.

Backing Up IoTDB Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-57 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.
Backup Policy	**Full backup at the first time and incremental backup subsequently**	Indicates a periodic backup policy. • Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times

Step 7 In **Configuration**, choose **IoTDB** > **IoTDB** under **Service data**.

Step 8 Select a backup directory type for **Path Type** of **IoTDB**.

The following backup directory types are supported:

RemoteHDFS: indicates that backup files are stored in the HDFS directory of the standby cluster.

If you select this option, set the following parameters:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Destination NameService Name**: indicates the NameService name of the standby cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
- **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
- **Destination Active NameNode IP Address**: indicates the service plane IP address of the active NameNode in the standby cluster.
- **Destination Standby NameNode IP Address**: indicates the service plane IP address of the standby NameNode in the standby cluster.
- **Destination NameNode RPC Port**: indicates the value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the standby cluster.
- **Target Path**: indicates the HDFS directory for storing standby cluster backup data. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.

Step 9 Set **Backup Content** to one or multiple service data records to be backed up.

You can select backup data using either of the following methods:

- Adding a backup data file
  a. Click **Add**.
  b. Select the table to be backed up under **File Directory**, and click **Add** to add the table to **Backup Content**.
  c. Click **OK**.
- Selecting using regular expressions
  a. Click **Query Regular Expression**.
  b. Enter the parent directory full path of the directory in the first text box as prompted. The directory must be the same as the existing directory, for example, **/root**.
  c. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to get all files or subdirectories in the parent directory, enter **([\s\S]*?)**. To get files whose names consist of letters and digits, for example, **file*1***, enter **file\d***.
  d. Enter a regular expression in the second text box. Standard regular expressions are supported. For example, to get objects containing **test**, enter **.*test.***. To get objects starting with **test**, enter **test.***. To get objects ending with **test**, enter **.*test**.
  e. Click **Refresh** to view the displayed directories in **Directory Name**.
  f. Click **Synchronize** to save the result.
  - **When entering regular expressions, click + or − to add or delete an expression.**
  - **If the selected table or directory is incorrect, click Clear Selected Node to deselect it.**
  - The backup directory must not contain files that have been written for a long time. Otherwise, the backup task will fail. Therefore, you are not

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

advised to perform operations on top-level directories such as **/user**, **/tmp**, and **/mr-history**.

Step 10 Click **Verify** to check whether the backup task is configured correctly.

The possible causes of the verification failure are as follows:

- The target NameNode IP address is incorrect.
- The data to be backed up does not exist.

Step 11 Click **OK**.

Step 12 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Data source_Task creation time*, and the subdirectory is used to save latest data source backup files. All the backup file sets are stored in the related snapshot directories.

----End

9.6.5.16 Backing Up Kafka Metadata

Scenarios

To ensure Kafka metadata security or before a major operation on ZooKeeper (such as upgrade or migration), you need to back up Kafka metadata. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

Kafka metadata backup tasks can be created on FusionInsight Manager. The system supports both automatic and manual backup modes.

MRS clusters support multiple data path types for backing up Kafka metadata.

- **LocalDir**: indicates that backup files are stored on the local disk of the active management node, with the standby management node automatically synchronizing the backup files.

  The default storage directory is *Data storage path*/**LocalBackup/**, for example, **/srv/BigData/LocalBackup**.
- **LocalHDFS**: indicates that backup files are stored in the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that backup files are stored in the HDFS directory of the standby cluster.
- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.
- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.
- **OBS**: indicates that backup files are stored in OBS.

Notes and Constraints

MRS 3.1.0 and later versions support the backup of Kafka metadata to OBS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- To back up data to a remote HDFS, the following conditions must be met:
  - A standby cluster for backing up data has been created. Its authentication mode must be the same as that of the active cluster.
  - If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The backup type, period, policy, and other specifications have been planned based on the service requirements and you have checked whether *Data storage path*/**LocalBackup/** has sufficient space on the active and standby management nodes.
- If you want to back up data to NAS, you have deployed the NAS server in advance.
- If you want to back up data to OBS, you have connected the current cluster to OBS and have the permission to access OBS.

Backing Up Kafka Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 Click **Create**.

Step 4 Set **Name** to the name of the backup task.

Step 5 Select the cluster to be operated from **Backup Object**.

Step 6 Set **Mode** to the type of the backup task.

Periodic indicates that the backup task is executed by the system periodically. **Manual** indicates that the backup task is executed manually.

Table 9-58 Periodic backup parameters

Parameter	Example Value	Description
Started	**2024/11/22 10:29:11**	Time when the task should start for the first time.
Period	**1 day**	Task execution interval. Supported values are **Hours** and **Days**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Backup Policy	**Full backup every time**	• Full backup at the first time and incremental backup subsequently • Full backup every time • Full backup once every *n* times Incremental backup is not supported for Manager data and component metadata. Only **Full backup every time** is supported. If **Path Type** is set to **NFS** or **CIFS**, incremental backup cannot be used. When incremental backup is used for NFS or CIFS backup, the latest full backup data is updated each time the incremental backup is performed. Therefore, no new recovery point is generated.

Step 7 In **Configuration**, select **Kafka**.

Step 8 Select a backup directory type for **Path Type** of **Kafka**.

Table 9-59 Path of backup data

Path Type	Parameter	Description
LocalDir	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
LocalHDFS	Target Path	HDFS directory for storing the backup files. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Target NameService Name	NameService name of the backup directory. The default value is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
RemoteHDFS	Destination NameService Name	Name of the NameService in the standby cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Target NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Target Path	HDFS directory in the standby cluster for storing backup data. • The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory. • The path cannot be configured as a default system directory, such as **/hbase** or **/user/hbase/backup**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Server Shared Path	Shared directory configured on the NAS server by the user. The shared directory of the server cannot be set to the root directory. The user group and owner group of the shared directory must be configured as **nobody:nobody**.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.
OBS (available in MRS 3.1.0 and later)	Target Path	OBS directory for storing backup data.
	Maximum Number of Backup Copies	Number of backup files that can be retained in the backup directory.

Step 9 Click **OK**.

Step 10 In the **Operation** column of the created task in the backup task list, click **More** and select **Back Up Now** to execute the backup task.

After the backup task is executed, the system automatically creates a subdirectory for each backup task in the backup directory. The format of the subdirectory name is *Backup task name_Task creation time*, and the subdirectory is used to save data source backup files. The format of the backup file name is *Version_Data source_Task execution time*.**tar.gz**.

----End

Helpful Links

Backup Kafka metadata may be restored if data is accidentally modified, deleted, or requires retrieval; if major ZooKeeper operations (such as upgrades or

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

significant data adjustments) cause exceptions in system data or fail to achieve the expected result; if all Kafka modules fail and become unavailable; or if data needs to be migrated to a new cluster. For detailed operations, refer to **Restoring Kafka Metadata**.

9.6.6 Restoring MRS Cluster Component Data

9.6.6.1 Restoring Manager Data (MRS 2.x and Earlier)

Scenarios

Metadata restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major metadata component operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

This section describes how to create metadata restoration tasks on MRS Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring Manager data.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **LocalHDFS**: indicates that data is restored from the HDFS directory of the current cluster.

NOTICE

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the data that is generated after the data backup and before the data restoration will be lost.
- Use the OMS data and LdapServer data backed up at the same time to restore data. Otherwise, the service and operation may fail.
- By default, MRS clusters use DBService to store Hive metadata.

Impact on the System

- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, the configurations of the components that depend on DBService may expire and these components need to be restarted.

Prerequisites

- The data in the OMS and LdapServer backup files has been backed up at the same time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The status of the OMS resources and the LdapServer instances is normal. If the status is abnormal, data restoration cannot be performed.
- The status of the cluster hosts and services is normal. If the status is abnormal, data restoration cannot be performed.
- The cluster host topologies during data restoration and data backup are the same. If the topologies are different, data restoration cannot be performed and you need to back up data again.
- The services added to the cluster during data restoration and data backup are the same. If the services are different, data restoration cannot be performed and you need to back up data again.
- The status of the active and standby DBService instances is normal. If the status is abnormal, data restoration cannot be performed.
- The upper-layer applications that depend on the MRS cluster have been stopped.
- On MRS Manager, you have stopped all the NameNode role instances whose data is to be recovered. Other HDFS role instances are running properly. After data is recovered, the NameNode role instances need to be restarted and cannot be accessed before the restart.
- You have checked whether NameNode backup files have been stored in the *Data save path***/LocalBackup/** directory on the active management node.

Restoring Manager Data

Step 1 Check the location of backup data.

1. On MRS Manager, choose **System** > **Back Up Data**.
2. In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records. In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:
   - **Backup Object**: indicates the backup data source.
   - **Backup Path**: indicates the full path where the backup files are stored.
3. Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 2 Create a restoration task.

1. On MRS Manager, choose **System** > **Recovery Management**.
2. On the page that is displayed, click **Create Restoration Task**.
3. Set **Task Name** to the name of the restoration task.

Step 3 Select restoration sources.

In **Configuration**, select the metadata component whose data is to be restored.

Step 4 Set the restoration parameters.

1. Select a backup directory type for **Path Type**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-60 Path for data restoration

Path Type	Parameter	Description
LocalDir	Source Path	Full path of the directory storing backup files. Path format: *Data storage path*/**LocalBackup**/*Backup task name_Task creation time*/*Data source_Task execution time*/*Version number_Data source_Task execution time*.**tar.gz**
LocalHDFS	Source Path	Full path of the HDFS directory storing backup files. Path format: *Backup path*/*Backup task name_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Source Instance Name	NameService name of the backup directory during restoration task execution. The default value is **hacluster**.

2. Click **OK**.

Step 5 Execute the restoration task.

In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and try to execute the task again by clicking **Start**.

Step 6 Determine what metadata has been restored.

- If the OMS and LdapServer metadata is restored, go to **Step 7**.
- If DBService data is restored, no further action is required.
- Restore NameNode data. On MRS Manager, choose **Services > HDFS > More > Restart Service**. The task is complete.

Step 7 Restarting Manager for the recovered data to take effect

1. In MRS Manager, Choose **LdapServer** > **More** > **Restart Service** and click **OK**. Wait until the LdapServer service is restarted successfully.
2. Log in to the active management node. For details, see **Checking MRS Active/Standby Management Nodes**.
3. Run the following command to restart OMS:

   sh ${BIGDATA_HOME}/om-0.0.1/sbin/restart-oms.sh

   The command is successfully executed if the following information is displayed:

   ```
   start HA successfully.
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. On MRS Manager, choose **KrbServer > More > Synchronize Configuration**. Do not select Restart the services and instances whose configuration has expired. Click **OK** and wait until the KrbServer service configuration is synchronized and restarted successfully.
5. Choose **Services > More > Synchronize Configuration**. Do not select Restart the services and instances whose configuration has expired. Click **OK** and wait until the cluster is configured and synchronized successfully.
6. Choose **Services > More > Stop Cluster**. After the cluster is stopped, choose **Services > More> Start Cluster**.

----End

9.6.6.2 Restoring Manager Data (MRS 3.x and Later Versions)

Scenarios

Manager data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major data adjustments on FusionInsight Manager cause exceptions in system data or fail to achieve the expected result; and when all modules fail and become unavailable.

Manager data restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring Manager data.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **LocalHDFS**: indicates that data is restored from the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.
- OBS: indicates that data is restored from OBS.

NOTE

To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the Manager data that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- MRS 3.1.0 and later versions support the storage of Manager data backup files to OBS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- In the restoration process, the Controller needs to be restarted and FusionInsight Manager cannot be logged in or operated during the restart.
- In the restoration process, all clusters need to be restarted and cannot be accessed during the restart.
- After data restoration, the data, such as system configuration, user information, alarm information, and audit information, that is generated after the data backup and before the data restoration will be lost. This may result in data query failure or cluster access failure.
- Once Manager data is restored, the system automatically forces each cluster's LDAP server to synchronize data from OLDAP.

Prerequisites

- To restore data from a remote HDFS, the following conditions must be met:
  - A standby cluster has been created for data restoration, and its data has been backed up. For details, see **Backing Up Manager Data (MRS 3.x and Later Versions)**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The status of the OMS resources and the LdapServer instances of each cluster is normal. If the status is abnormal, data restoration cannot be performed.
- The status of the cluster hosts and services is normal. If the status is abnormal, data restoration cannot be performed.
- The cluster host topologies during data restoration and data backup are the same. If the topologies are different, data restoration cannot be performed and you need to back up data again.
- The services added to the cluster during data restoration and data backup are the same. If the topologies are different, data restoration cannot be performed and you need to back up data again.
- The upper-layer applications that depend on the cluster are stopped.

Restoring Manager Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 Choose **O&M** > **Backup and Restoration** > **Restoration Management**. On the displayed page, click **Create**.

Step 5 Set **Task Name** to the name of the restoration task.

Step 6 Set **Recovery Object** to **OMS**.

Step 7 Select **OMS**.

Step 8 Select a backup directory type for **Path Type** of **OMS**.

Table 9-61 Path for data restoration

Type	Parameter	Description
LocalDir	Source Path	Backup file to be restored. File name format: *Version_Data source_Task execution time*.tar.gz
LocalHDFS	Source Path	Full path of the HDFS directory storing backup files. Path format: *Backup path*/*Backup task name_Task creation time*/*Version_Data source_Task execution time*.tar.gz
	Cluster for Restoration	Name of the cluster used to execute the restoration task.
	Source NameService Name	NameService name of the backup directory during restoration task execution. The default value is **hacluster**.
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Source Cluster	Cluster containing the YARN queue used by the data you want to restore.
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Source Path	Full path of the backup server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
OBS (available in MRS 3.1.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

Step 9 Click **OK**.

Step 10 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

Step 11 Log in to the active and standby management nodes as user **omm**.

Step 12 Run the following command to restart OMS:

sh ${BIGDATA_HOME}/om-server/om/sbin/restart-oms.sh

The command is successfully executed if the following information is displayed:

```
start HA successfully.
```

Run **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** to check whether **HAAllResOK** of the management node is **Normal** and whether FusionInsight Manager can be logged in again. If yes, OMS is restarted successfully.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 On FusionInsight Manager, click **Cluster**, click the name of the target cluster, and choose **Services** > **KrbServer**. On the displayed page, choose **More** > **Synchronize Configuration**, click **OK**, and wait for the KrbServer configuration to be synchronized and the service to be restarted.

Step 14 Choose **Cluster**, click the name of the desired cluster, and choose **More** > **Synchronize Configurations**, click **OK**, and wait until the cluster configuration is synchronized successfully.

Step 15 On FusionInsight Manager, click **Cluster**, click the name of the target cluster, and choose **More** > **Restart**. On the displayed page, enter the password of the current login user, click **OK**, and wait for the cluster to be restarted.

----End

9.6.6.3 Restoring CDL Service Data

Scenarios

CDL data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major CDL operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

CDL metadata is stored in DBService and Kafka. You can create DBService and Kafka restoration tasks on FusionInsight Manager to restore CDL service data. The system supports manual data restoration only.

NOTICE

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the DBService and Kafka data that is generated after the data backup and before the data restoration will be lost.
- By default, MRS clusters use DBService to store metadata of Hive, Hue, Loader, Spark, CDL, and Oozie. Restoring DBService data will restore the metadata of all these components.

Impact on the System

- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, the configurations of the components that depend on DBService may expire and these components need to be restarted.
- After the metadata is restored, the offset information stored on ZooKeeper by Kafka consumers is rolled back, resulting in repeated consumption.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- To restore data from a remote HDFS, the following conditions must be met:
  - A standby cluster has been created for data restoration, and its data has been backed up. For details, see **Backing Up CDL Service Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The status of the active and standby DBService instances is normal. If the status is abnormal, data restoration cannot be performed.
- The Kafka service has been stopped. After the restoration is complete, start the Kafka service.

Restoring CDL Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In the **Restoration Configuration** area, select **DBService** and **Kafka**.

Step 9 Select a backup directory type for **Path Type** of **DBService**. For details about how to configure the parameters, see **Step 9**.

Step 10 Select a backup directory type for **Path Type** of **Kafka**. For details about how to configure the parameters, see **Step 9**.

Step 11 Click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.4 Restoring ClickHouse Metadata

Scenarios

ClickHouse metadata restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major ClickHouse operations (such as upgrades or migration) cause exceptions in system data or fail to achieve the expected result; when all ClickHouse modules fail and become unavailable; and when data is migrated to a new cluster.

ClickHouse metadata restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring ClickHouse metadata.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- OBS: indicates that data is restored from OBS.

NOTE

To restore ClickHouse metadata when the service is running properly, it is recommended that you manually back up the latest ClickHouse metadata prior to restoration. Otherwise, the ClickHouse metadata that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- MRS 3.1.0 or later supports this function.
- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- ClickHouse metadata restoration and service data restoration cannot be performed simultaneously. Otherwise, service data restoration fails. You are advised to restore service data only after metadata restoration is complete.
- MRS 3.3.0 and later support data restoration from OBS.

Impact on the System

- After the metadata is restored, the data generated after the data backup and before the data restoration is lost.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After the metadata is restored, the ClickHouse upper-layer applications need to be started.

Prerequisites

- You have checked the path for storing ClickHouse metadata backup files.
- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up ClickHouse Metadata**. If the active and standby clusters are deployed in security mode and they are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active and standby clusters are deployed in normal mode, mutual trust is not required.
- In an active/standby cluster, the value of **HADOOP_RPC_PROTECTION** of ClickHouse must be the same as that of **hadoop.rpc.protection** in the HDFS when you restore data from the remote HDFS to the local host.
- To restore backup data from another MRS ClickHouse cluster to this cluster, the following requirements must be met:
  a. The two clusters are of the same MRS version.
  b. The two clusters are in the same mode.
  c. The two clusters share the same ClickHouse topology, including shards and replicas.

Restoring ClickHouse Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the **Operation** column of the specified task in the task list, choose **More** > **View History**.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **ClickHouse** under **Metadata and other data**.

Step 9 Select a backup directory type for **Path Type** of **ClickHouse**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-62 Path for data restoration

Path Type	Parameter	Description
LocalDir	Source Path	Name of the backup file to be restored. To obtain the file name, log in to the active OMS node, navigate to the backup path copied in **Step 3**, and record the metadata package name. The file name follows the format: *Backup task name_Data source_Task execution time***.tar.gz**.
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source Active NameNode IP Address	Service plane IP address of the active NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the active NameNode. This parameter is available only for MRS 3.2.0 and later clusters.
	Source Standby NameNode IP Address	Service plane IP address of the standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the standby NameNode. This parameter is available only for MRS 3.2.0 and later clusters.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the NameNode. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Source NameNode RPC Port	Value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the destination cluster. This parameter is available only for MRS 3.2.0 and later clusters.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. For details, see **Backup Path** obtained in **Step 3**. Path format: *Backup path*/*Backup task name*_*Data source*_*Task creation time*
OBS	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name*_*Data source*_*Task creation time*/*Version*_*Data source*_*Task execution time*.tar.gz

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row where the created task is located, and click **Start** in the **Operation** column. In the displayed dialog box, click **OK** to start the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

Step 12 Choose **Cluster** > **Services** and start the ClickHouse service.

----End

9.6.6.5 Restoring ClickHouse Service Data

Scenarios

ClickHouse data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major ClickHouse operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

ClickHouse service data restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

The ClickHouse backup and restoration functions cannot identify the service and structure relationships of objects such as ClickHouse tables, indexes, and views. When executing backup and restoration tasks, you need to manage unified restoration points based on service scenarios to ensure proper service running.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

MRS clusters support multiple data path types for restoring ClickHouse service data.

- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- OBS: indicates that data is restored from OBS.

NOTE

To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the ClickHouse data that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- MRS 3.1.0 or later supports this function.
- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- ClickHouse metadata restoration and service data restoration cannot be performed simultaneously. Otherwise, service data restoration fails. You are advised to restore service data only after metadata restoration is complete.
- MRS 3.3.0-LTS.1 and later versions support the storage of ClickHouse service data backup files to OBS.

Impact on the System

- During data restoration, user authentication stops and users cannot create new connections.
- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, the ClickHouse upper-layer applications need to be started.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up ClickHouse Service Data**. If the active and standby clusters are deployed in security mode and they are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active and standby clusters are deployed in normal mode, mutual trust is not required.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The database for storing restored data tables, the HDFS save path of data tables, and the list of users who can access restored data are planned.
- The ClickHouse backup file save path is correct.
- The ClickHouse upper-layer applications are stopped.
- In an active/standby cluster, the value of **HADOOP_RPC_PROTECTION** of ClickHouse must be the same as that of **hadoop.rpc.protection** in the HDFS when you restore data from the remote HDFS to the local host.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the cluster to be restored does not contain the metadata of the backup service data, you must restore the corresponding backup metadata before restoring the service data, or restore both the backup metadata and backup service data simultaneously.
- To restore backup data from another MRS ClickHouse cluster to this cluster, the following requirements must be met:
  a. The two clusters are of the same MRS version.
  b. The two clusters are in the same mode.
  c. The two clusters share the same ClickHouse topology, including shards and replicas.
  d. The number of ClickHouse disk partitions and disk capacity of the cluster to be restored must be greater than or equal to those of the backup cluster.

Restoring ClickHouse Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **ClickHouse** under **Service data**.

Step 9 Select a backup directory type for **Path Type** of **ClickHouse**.

Table 9-63 Path for data restoration

Path Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source Active NameNode IP Address	Service plane IP address of the active NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the active NameNode. This parameter is available for clusters of MRS 3.2.0 or later.
	Source Standby NameNode IP Address	Service plane IP address of the standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the standby NameNode. This parameter is available for clusters of MRS 3.2.0 or later.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster. Log in to MRS Manager of the standby cluster, choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and check the service plane IP address of the NameNode. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
	Source NameNode RPC Port	Value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the destination cluster. This parameter is available for clusters of MRS 3.2.0 or later.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. For details, see **Backup Path** obtained in **Step 3**. Path format: *Backup path*/*Backup task name_Data source_Task creation time*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**. This parameter is available only for MRS 3.1.0 and MRS 3.1.2 clusters.
OBS	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name*_*Data source*_*Task creation time*/*Version*_*Data source*_*Task execution time*.tar.gz

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.6 Restoring DBService Metadata

Scenarios

DBService data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major DBService operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

DBService metadata restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring DBService metadata.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **LocalHDFS**: indicates that data is restored from the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.
- OBS: indicates that data is restored from OBS.

NOTICE

- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the DBService data that is generated after the data backup and before the data restoration will be lost.
- By default, MRS clusters use DBService to store metadata of Hive, Hue, Loader, Spark, CDL, and Oozie. Restoring DBService data will restore the metadata of all these components.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- MRS 3.1.0 and later support data restoration from OBS.

Impact on the System

- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, the configurations of the components that depend on DBService may expire and these components need to be restarted.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up DBService Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The status of the active and standby DBService instances is normal. If the status is abnormal, data restoration cannot be performed.

Restoring DBService Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In the **Restoration Configuration** area, select **DBService**.

Step 9 Select a backup directory type for **Path Type** of **DBService**.

Table 9-64 Path for data restoration

Type	Parameter	Description
LocalDir	Source Path	Backup file to be restored. File name format: *Version_Data source_Task execution time*.tar.gz
LocalHDFS	Source Path	Full path of the HDFS directory storing backup files. Path format: *Backup path*/*Backup task name_Task creation time*/*Version_Data source_Task execution time*.tar.gz
	Source NameService Name	NameService name of the backup directory during restoration task execution. The default value is **hacluster**.
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster, or to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Source Path	Full path of the backup server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
OBS (available in MRS 3.1.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.7 Restoring Doris Service Data

Scenarios

Doris data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major Doris operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Doris service data restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

When executing backup and restoration tasks, you need to manage unified restoration points based on service scenarios to ensure proper service running.

MRS clusters support multiple data path types for restoring Doris service data.

- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- OBS: indicates that data is restored from OBS.

NOTE

To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the Doris data that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- This section applies only to MRS 3.5.0 or later clusters.
- Data restoration can be performed only when the system version is consistent with the version used during data backup.

Impact on the System

After the data is restored, the data generated after the data backup and before the data restoration is lost.

Prerequisites

- To restore data from a remote HDFS, the following conditions must be met:
  - A standby cluster has been created for data restoration, and its data has been backed up. For details, see **Backing Up Doris Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
  - At least one DBroker instance of the Doris service has been deployed in the active cluster.
  - Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
  - The value of **hadoop.rpc.protection** of Doris must be the same as that of **hadoop.rpc.protection** of HDFS in both active and standby clusters.
- If you want to restore data from OBS, you have connected the Doris cluster to OBS and have the permission to access OBS.
- The database for storing restored data tables, the location for storing the data tables in HDFS, and the list of users who can access the restored data have been planned.
- Check the path for storing Doris backup files.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Stop the upper-layer Doris applications.

Restoring Doris Service Data

Step 1 Log in to MRS Manager and choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 2 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 3 Choose **Restoration Management** and click **Create**.

Step 4 Set **Task Name** to the name of the restoration task.

Step 5 Select the cluster to be operated from **Recovery Object**.

Step 6 In **Restoration Configuration**, select **Doris** under **Service data**.

Step 7 Select a backup directory type for **Path Type** of **Doris**.

Table 9-65 Path for data restoration

Path Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. You can obtain it from the **NameService Management** page of HDFS in the standby cluster. For example, the name is **hacluster**.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the NameNode in the standby cluster.
	Source NameNode RPC Port	Value of **dfs.namenode.rpc.port** in the HDFS configuration of the standby cluster.
	DBroker IP	Service plane IP address of the DBroker role in the cluster. The DBroker is used to transmit data during restoration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Path Type	Parameter	Description
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. For details, see **Backup Path** obtained in **Step 2**. Path format: *Backup path*/*Backup task name_Data source_Task creation time*
OBS	Source Path	Full path of the OBS directory storing backup files. Specify this path by referring to **Step 2**. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/

Step 8 Click **Refresh** and select a Doris backup file set that has been backed up.

Step 9 In the **Data Configuration** area, select one or more pieces of backup data for **Select Data** based on service requirements.

Configuration restrictions are as follows:

- There is a database with the same name as the original database of the selected backup data in the Doris of the cluster.
- The backup data is restored to the backup table with the same name as the original table in the database.
- If there is a table with the same name in Doris, ensure that the structures of the two tables are the same, including table names, columns, partitions, and materialized views.

Step 10 Set **Original Configurations** to **true**, indicating that the configuration of the backup data, such as the number of copies, will be used. If this parameter is set to **false**, the default configuration is used to create a table.

Step 11 Click **OK**.

Step 12 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.8 Restoring Flink Metadata

Scenarios

Flink metadata restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major Flink operations (such as upgrades or migration) cause exceptions in system data or fail

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

to achieve the expected result; when all Flink modules fail and become unavailable; and when data is migrated to a new cluster.

Flink metadata restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring Flink metadata.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **LocalHDFS**: indicates that data is restored from the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- OBS: indicates that data is restored from OBS.

NOTE

To restore Flink metadata when the service is running properly, it is recommended that you manually back up the latest Flink metadata prior to restoration. Otherwise, the Flink metadata that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- Flink metadata restoration and service data restoration cannot be performed simultaneously. Otherwise, service data restoration fails. You are advised to restore service data only after metadata restoration is complete.
- MRS 3.5.0 and later support data restoration from OBS.

Impact on the System

- Before restoring the metadata, you need to stop the Flink service. During this period, all upper-layer applications are affected and cannot work properly.
- After the metadata is restored, the data generated after the data backup and before the data restoration is lost.
- After the metadata is restored, the Flink upper-layer applications of Solr need to be started.

Prerequisites

- You have checked the path for storing Flink metadata backup files.
- The Flink service has been stopped before its metadata is restored.
- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up Flink Metadata**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.

Restoring Flink Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the **Operation** column of the specified task in the task list, choose **More** > **View History**.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **Flink** under **Metadata and other data**.

Step 9 Select a backup directory type for **Path Type** of **Flink**.

Table 9-66 Path for data restoration

Type	Parameter	Description
LocalDir	Source Path	Backup file to be restored. File name format: *Version_Data source_Task execution time*.tar.gz
LocalHDFS	Source Path	Full path of the HDFS directory storing backup files. Path format: *Backup path*/*Backup task name_Task creation time*/*Version_Data source_Task execution time*.tar.gz
	Source NameService Name	NameService name of the backup directory during restoration task execution. The default value is **hacluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
OBS (available in MRS 3.5.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row where the created task is located, and click **Start** in the **Operation** column. In the displayed dialog box, click **OK** to start the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

Step 12 Choose **Cluster** > **Services** and start the Flink service.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.6.6.9 Restoring HBase Metadata

Scenarios

To ensure HBase metadata security (including tableinfo files and HFiles) or before a major operation on HBase system tables (such as upgrade or migration), you need to back up HBase metadata to prevent HBase service unavailability caused by HBase system table directory or file damages. The backup data can be used to recover the system if an exception occurs or the operation has not achieved the expected result, minimizing the adverse impacts on services.

HBase metadata restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring HBase metadata.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.
- OBS: indicates that data is restored from OBS.

NOTICE

- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the HBase data that is generated after the data backup and before the data restoration will be lost.
- It is recommended that each data restoration task restore the metadata of only one component to avoid affecting the restoration of other components when a service or instance is stopped. If data of multiple components is restored at the same time, data restoration may fail.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- HBase metadata cannot be restored concurrently with NameNode metadata. Attempting both at once will cause restoration to fail.
- MRS 3.1.0 and later support data restoration from OBS.

Impact on the System

- Before restoring the metadata, you need to stop the HBase service, during which the HBase upper-layer applications are unavailable.
- After the metadata is restored, the data generated after the data backup and before the data restoration is lost.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After the metadata is restored, you need to start the upper-layer applications of HBase.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up HBase Metadata**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- You have checked the path for storing HBase metadata backup files.
- The HBase service has been stopped before its metadata is restored.

Restoring HBase Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **HBase** under **Metadata and other data**.

Step 9 Select a backup directory type for **Path Type** of **HBase**.

Table 9-67 Path for data restoration

Type	Parameter	Description
LocalDir	Source Path	Backup file to be restored. File name format: *Version_Data source_Task execution time*.tar.gz

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*Backup task name_Data source_Task creation time*|*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*|*Backup task name_Data source_Task creation time*|*Version_Data source_Task execution time*.**tar.gz**
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name*_*Data source*_*Task creation time*/*Version*_*Data source*_*Task execution time*.**tar.gz**
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Source Path	Full path of the backup server directory storing backup files. Path format: *Backup path*/*Backup task name*_*Data source*_*Task creation time*/*Version*_*Data source*_*Task execution time*.**tar.gz**
OBS (available in MRS 3.1.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name*_*Data source*_*Task creation time*/*Version*_*Data source*_*Task execution time*.**tar.gz**

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.6.6.10 Restoring HBase Service Data

Scenarios

HBase service data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major HBase operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

HBase service data restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring HBase service data.

- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.

NOTICE

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the HBase data that is generated after the data backup and before the data restoration will be lost.

Impact on the System

- During data restoration, the system disables the HBase table being restored, and the table cannot be accessed at that time. The data restoration process takes several minutes, during which the HBase upper-layer applications are unavailable.
- During data restoration, user authentication stops and users cannot create new connections.
- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, you need to start the upper-layer applications of HBase.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up HBase Service Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Between MRS Clusters. If the active cluster is deployed in normal mode, mutual trust is not required.

- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The directory for saving the backup file has been checked.
- The HBase upper-layer applications have been stopped.

Restoring HBase Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **HBase** under **Service data**.

Step 9 Select a backup directory type for **Path Type** of **HBase**.

Table 9-68 Path for data restoration

Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*xxx*|*Backup task name_Data source_Task creation time*
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*xxx*|*Backup task name_Data source_Task creation time*
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*xxx*|*Backup task name_Data source_Task creation time*
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*xxx*|*Backup task name_Data source_Task creation time*
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

Step 10 Set **Backup Data** column in **Data Configuration** to one or multiple backup data sources to be recovered. In the **Target Namespace** column, specify the target naming space after backup data recovery.

You are advised to set **Target Namespace** to a location that is different from the backup naming space.

Step 11 Set **Force recovery** to **true** to forcibly restore all backup data when a table with the same name already exists. Any data added to the table after the backup will be lost during restoration. If you set the parameter to **false**, the restoration task is not executed if a data table with the same name exists.

Step 12 Click **Verify** to check whether the restoration task is configured correctly.

- If the queue name is incorrect, the verification fails.
- If the specified naming space does not exist, the verification fails.
- If the forcible overwrite conditions are not met, the verification fails.

Step 13 Click **OK** to save the settings.

Step 14 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

Step 15 Check whether HBase data is restored in an environment where HBase is newly installed or reinstalled.

- If yes, the administrator needs to set new permission for roles on FusionInsight Manager based on the original service plan.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, no further operation is required.

----End

9.6.6.11 Restoring HDFS NameNode Metadata

Scenarios

NameNode data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major NameNode operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

NameNode metadata restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring NameNode metadata.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.
- OBS: indicates that data is restored from OBS.

NOTICE

- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the NameNode data that is generated after the data backup and before the data recovery will be lost.
- It is recommended that each data restoration task restore the metadata of only one component to avoid affecting the restoration of other components when a service or instance is stopped. If data of multiple components is restored at the same time, data restoration may fail.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- HBase metadata cannot be restored concurrently with NameNode metadata. Attempting both at once will cause restoration to fail.
- MRS 3.1.0 and later support data restoration from OBS.

Impact on the System

- After the data is restored, the data generated after the data backup and before the data restoration is lost.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After the data is recovered, the NameNode needs to be restarted and is unavailable during the restart.
- After data is restored, metadata and service data may not be matched, the HDFS enters the security mode, and the HDFS service fails to be started. .

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up HDFS NameNode Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- On FusionInsight Manager, all the NameNode role instances whose data is to be recovered are stopped. Other HDFS role instances must keep running. After data is recovered, the NameNode role instances need to be restarted. The NameNode role instances cannot be accessed during the restart.
- The NameNode backup files are stored ***Data path*/LocalBackup/** on the active management node.

Restoring HDFS NameNode Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **Cluster** > **Services** > **HDFS**. On the displayed page, click **Instances** and then **NameNode** to check whether the NameNode instances whose data is to be restored are stopped. If they are not, click **Stop Instance** to stop them.

Step 3 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 4 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 5 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 6 Click **Create**.

Step 7 Set **Task Name** to the name of the restoration task.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Select the cluster to be operated from **Recovery Object**.

Step 9 In the **Restoration Configuration** area, select **NameNode**.

Step 10 Select a backup directory type for **Path Type** of **NameNode**.

Table 9-69 Path for data restoration

Type	Parameter	Description
LocalDir	Source Path	Backup file to be restored. File name format: *Version_Data source_Task execution time*.tar.gz
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*Backup task name_Data source_Task creation time*|*Version_Data source_Task execution time*.tar.gz
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Password	Password used to connect to the server over SFTP.
	Source Path	Full path of the backup server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
OBS	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.

Step 11 Click **OK**.

Step 12 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

Step 13 On FusionInsight Manager, click **Cluster**, click the name of the desired cluster, and choose **Services** > **HDFS**. On the displayed page, click **Configurations** and click **All Configurations**.

On the displayed page, enter the password of the administrator who has logged in for authentication and click **OK**. After the system displays "Operation succeeded", click **Finish**. The service is started successfully.

----End

9.6.6.12 Restoring HDFS Service Data

Scenarios

HDFS data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major HDFS operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

HDFS service data restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring HDFS service data.

- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.
- OBS: indicates that data is restored from OBS.

NOTE

To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the HDFS data that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- The HDFS restoration operation cannot be performed on directories used for running YARN tasks (for example, **/tmp/logs**, **/tmp/archived**, and **/tmp/hadoop-yarn/staging**). Attempting restoration on these directories will cause DistCp tasks to fail due to file loss.
- MRS 3.5.0 and later support data restoration from OBS.

Impact on the System

- During data restoration, user authentication stops and users cannot create new connections.
- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is recovered, the HDFS upper-layer applications need to be started.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up HDFS Service Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The HDFS backup file save path is correct.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The HDFS upper-layer applications are stopped.

Restoring HDFS Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **HDFS** under **Service data**.

Step 9 Select a backup directory type for **Path Type** of **HDFS**.

Table 9-70 Path for data restoration

Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*|*Backup task name_Data source_Task creation time*|*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Source Path	Full path of the backup server directory storing backup files. Path format: *Backup path*/*Backup task name*_Data source_*Task creation time*/*Version*_Data source_*Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
OBS (available in MRS 3.5.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name*_Data source_*Task creation time*/*Version*_Data source_*Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a backed-up HDFS directory in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 In the **Backup Data** column of the **Data Configuration** page, select one or more pieces of backup data that needs to be restored based on service requirements. In the **Target Path** column, specify the target location after backup data restoration.

You are advised to set **Target Path** to a new path that is different from the backup path.

Step 11 Click **Verify** to check whether the restoration task is configured correctly.

- If the queue name is incorrect, the verification fails.
- If the specified directory to be restored does not exist, the verification fails.

Step 12 Click **OK**.

Step 13 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.13 Restoring Hive Service Data

Scenarios

Hive data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major Hive operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

Hive service data restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring Hive service data.

- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- SFTP: indicates that data is restored from the server using the SFTP protocol.
- OBS: indicates that data is restored from OBS.

Hive backup and restoration cannot identify the service and structure relationships of objects such as Hive tables, indexes, and views. When executing backup and restoration tasks, you need to manage unified restoration points based on service scenarios to ensure proper service running.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the Hive data that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

Data restoration can be performed only when the system version is consistent with the version used during data backup.

- MRS 3.5.0 and later support data restoration from OBS.

Impact on the System

- During data restoration, user authentication stops and users cannot create new connections.
- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, you need to start the upper-layer applications of Hive.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up Hive Service Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The database for storing restored data tables, the HDFS save path of data tables, and the list of users who can access restored data are planned.
- The Hive backup file save path is correct.
- The Hive upper-layer applications are stopped.

Restoring Hive Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In the **Restoration Configuration** area, select **Hive**.

Step 9 Select a backup directory type for **Path Type** of **Hive**.

Table 9-71 Path for data restoration

Type	Parameter	Description
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*|*Backup task name_Data source_Task creation time*|*Version_Data source_Task execution time*.tar.gz
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a Hive backup file set that has been backed up in the standby cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*|*Backup task name*_*Data source*_*Task creation time*|*Version*_*Data source*_*Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a Hive backup file set that has been backed up in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a Hive backup file set that has been backed up in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
SFTP	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the server where the backup data is stored.
	Port	Port number used by the SFTP protocol to connect to the backup server. The default value is **22**.
	Username	Username used to connect to the server over SFTP.
	Password	Password used to connect to the server over SFTP.
	Source Path	Full path of the backup server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select a Hive backup file set that has been backed up in the standby cluster.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.
OBS (available in MRS 3.5.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
	Restoration Point List	Click **Refresh** and select an OBS directory that has already been backed up.
	Target NameService Name	Name of the target NameService for the backup directory. The default value is **hacluster**.
	Maximum Number of Maps	Maximum number of maps in a MapReduce task. The default value is **20**.
	Maximum Map Bandwidth (MB/s)	Maximum bandwidth of a map. The default value is **100**.

Step 10 Set **Backup Data** in the **Data Configuration** to one or multiple backup data sources to be recovered based on service requirements. In the **Target Database** and **Target Path** columns, specify the target database and file save path after backup data recovery.

Configuration restrictions:

- Data can be restored to the original database, but data tables must be stored in a new path that is different from the backup path.
- To restore Hive index tables, select the Hive data tables that correspond to the Hive index tables to be restored.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If a new restoration directory is selected to avoid affecting the current data, HDFS permission must be manually granted so that users who have permission of backup tables can access this directory.
- Data can be restored to other databases. In this case, HDFS permission must be manually granted so that users who have permission of backup tables can access the HDFS directory that corresponds to the database.

Step 11 Set **Force recovery** to **true** to forcibly restore all backup data when a table with the same name already exists. Any data added to the table after the backup will be lost during restoration. If you set the parameter to **false**, the restoration task is not executed if a data table with the same name exists.

Step 12 Click **Verify** to check whether the restoration task is configured correctly.

- If the queue name is incorrect, the verification fails.
- If the specified directory to be restored does not exist, the verification fails.
- If the forcible overwrite conditions are not met, the verification fails.

Step 13 Click **OK**.

Step 14 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.14 Restoring IoTDB Metadata

Scenarios

To ensure IoTDB metadata security and prevent the IoTDB service from being unavailable due to IoTDB file damage, IoTDB metadata needs to be backed up. In this way, the system can restore data timely when an exception is reported or an operation does not achieve the expected result, minimizing the impact on services.

System administrators can create an IoTDB restoration task on FusionInsight Manager. Only manual restoration tasks are supported.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the IoTDB data that is generated after the data backup and before the data restoration will be lost.
- It is recommended that each data restoration task restore the metadata of only one component to avoid affecting the restoration of other components when a service or instance is stopped. If data of multiple components is restored at the same time, data restoration may fail.

Impact on the System

After the metadata is restored, the data generated after the data backup and before the data restoration is lost.

Restoring IoTDB Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, select **IoTDB** under **Metadata and other data**.

Step 9 Select a backup directory type for **Path Type**.

The configurations vary based on backup directory types:

- **LocalDir**: indicates that the backup files are stored on the local disk of the active management node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If you select this option, you also need to set **Source Path**, which indicates the backup file to be restored, for example, *Version_Data source_Task execution time*.**tar.gz**.

- **NFS**: indicates that backup files are stored in NAS using the NFS protocol.

  If you select this option, configure the following parameters:

  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
  - **Server IP Address**: indicates the IP address of the NAS server.
  - **Source Path**: indicates the complete path of the backup file on the NAS server, for example, *Backup path/Backup task name_Data source_Task creation time/Version_Data source_Task execution time*.**tar.gz**.

- **RemoteHDFS**: indicates that the backup files are stored in the HDFS directory of the standby cluster.

  If you select this option, configure the following parameters:

  - **Source NameService Name**: indicates the NameService name of the backup data cluster, for example, **hacluster**. You can obtain it from the **NameService Management** page of HDFS in the standby cluster.
  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
  - **Source Active NameNode IP Address**: indicates the service plane IP address of the active NameNode in the standby cluster.
  - **Source Standby NameNode IP Address**: indicates the service plane IP address of the standby NameNode in the standby cluster.
  - **Source NameNode RPC Port**: indicates the value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the standby cluster.
  - **Source Path**: indicates the full path of HDFS directory for storing backup data of the standby cluster, for example, *Backup path/Backup task name_Data source_Task creation time/Version_Data source_Task execution time*.**tar.gz**.

- **CIFS**: indicates that backup files are stored in NAS using the CIFS protocol.

  If you select this option, configure the following parameters:

  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
  - **Server IP Address**: indicates the IP address of the NAS server.
  - **Port**: indicates the port number used to connect to the NAS server over the CIFS protocol. The default value is **445**.
  - **Username**: indicates the username set when the CIFS protocol is configured.
  - **Password**: indicates the password set when the CIFS protocol is configured.
  - **Source Path**: indicates the complete path of the backup file on the NAS server, for example, *Backup path/Backup task name_Data source_Task creation time/Version_Data source_Task execution time*.**tar.gz**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **SFTP**: indicates that backup files are stored in the server using the SFTP protocol.

  If you select this option, configure the following parameters:

  - **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
  - **Server IP Address**: indicates the IP address of the server where the backup data is stored.
  - **Port**: indicates the port number used to connect to the backup server over the SFTP protocol. The default value is **22**.
  - **Username**: indicates the username for connecting to the server using the SFTP protocol.
  - **Password**: indicates the password for connecting to the server using the SFTP protocol.
  - **Source Path**: indicates the complete path of the backup file on the backup server, for example, *Backup path/Backup task name_Data source_Task creation time/Version_Data source_Task execution time*.**tar.gz**.

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

Step 12 Choose **Cluster** > **Services** and start the IoTDB service.

----End

9.6.6.15 Restoring IoTDB Service Data

Scenarios

IoTDB data restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major IoTDB operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

System administrators can create an IoTDB restoration task on FusionInsight Manager. Only manual restoration tasks are supported.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- To restore data when the service is running properly, it is recommended that you manually back up the latest management data before performing data restoration. Otherwise, the IoTDB data that is generated after the data backup and before the data restoration will be lost.

Impact on the System

- During data restoration, user authentication stops and users cannot create new connections.
- After the data is restored, the data generated after the data backup and before the data restoration is lost.
- After the data is restored, the IoTDB upper-layer applications need to be started.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up IoTDB Service Data**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The IoTDB backup file save path is correct.
- The IoTDB upper-layer applications are stopped.

Restoring IoTDB Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In **Restoration Configuration**, choose **IoTDB** > **IoTDB** under **Service Data**.

Step 9 Select a backup directory type for **Path Type** of **IoTDB**.

The following backup directory types are supported:

RemoteHDFS: indicates that the backup files are stored in the HDFS directory of the standby cluster.

If you select this option, configure the following parameters:

- **Source NameService Name**: indicates the NameService name of the backup data cluster, for example, **hacluster**. You can obtain it from the **NameService Management** page of HDFS in the standby cluster.
- **IP Mode**: indicates the IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
- **Source Active NameNode IP Address**: indicates the service plane IP address of the active NameNode in the standby cluster.
- **Source Standby NameNode IP Address**: indicates the service plane IP address of the standby NameNode in the standby cluster.
- **Source NameNode RPC Port**: indicates the value of **dfs.namenode.rpc.port** in the HDFS basic configuration of the standby cluster.
- **Source Path**: indicates the full path of HDFS directory for storing backup data of the standby cluster, for example, *Backup path/Backup task name_Data source_Task creation time*.
- **Recovery Point List**: Click **Refresh** and select an IoTDB directory that has been backed up in the standby cluster.

Step 10 In the **Backup Data** column of the **Data Configuration** page, select one or more pieces of backup data that needs to be restored based on service requirements. In the **Target Path** column, specify the target location after backup data restoration.

You are advised to set **Target Path** to a new path that is different from the backup path.

Step 11 Set **Force recovery** to **true** to forcibly restore all backup data when a table with the same name already exists. Any data added to the table after the backup will be lost during restoration. If you set the parameter to **false**, the restoration task is not executed if a data table with the same name exists.

Step 12 Click **Verify** to check whether the restoration task is configured correctly.

- If the queue name is incorrect, the verification fails.
- If the specified directory to be restored does not exist, the verification fails.

Step 13 Click **OK**.

Step 14 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

----End

9.6.6.16 Restoring Kafka Metadata

Scenarios

Kafka metadata restoration is required in the following scenarios: when data is unexpectedly modified or deleted and requires retrieval; when major ZooKeeper operations (such as upgrades or significant adjustments) cause exceptions in system data or fail to achieve the expected result; when all modules fail and become unavailable; and when data is migrated to a new cluster.

Kafka metadata restoration tasks can be created on FusionInsight Manager. The system supports manual data restoration only.

MRS clusters support multiple data path types for restoring Kafka metadata.

- **LocalDir**: indicates that data is restored from the local disk of the active management node.
- **LocalHDFS**: indicates that data is restored from the HDFS directory of the current cluster.
- **RemoteHDFS**: indicates that data is restored from the HDFS directory of the standby cluster.
- **NFS**: indicates that data is restored from the NAS using the NFS protocol.
- **CIFS**: indicates that data is restored from the NAS using the CIFS protocol.
- OBS: indicates that data is restored from OBS.

NOTE

To restore Kafka metadata when the service is running properly, it is recommended that you manually back up the latest Kafka metadata prior to restoration. Otherwise, the Kafka metadata that is generated after the data backup and before the data restoration will be lost.

Notes and Constraints

- Data restoration can be performed only when the system version is consistent with the version used during data backup.
- MRS 3.1.0 and later support data restoration from OBS.

Impact on the System

- After the metadata is restored, the data generated after the data backup and before the data restoration is lost.
- After the metadata is restored, the offset information stored on ZooKeeper by Kafka consumers is rolled back, resulting in repeated consumption.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- If you need to restore data from a remote HDFS, a standby cluster has been created and the data has been backed up. For details, see **Backing Up Kafka Metadata**. If the active cluster is deployed in security mode and the active and standby clusters are not managed by the same FusionInsight Manager, mutual trust has been configured. For details, see **Configuring Mutual Trust Between MRS Clusters**. If the active cluster is deployed in normal mode, mutual trust is not required.
- Cross-cluster replication has been configured for the active and standby clusters. For details, see **Enabling MRS Inter-Cluster Replication**.
- Time is consistent between the active and standby clusters, with the NTP services on both clusters configured to use the same time source.
- The Kafka service is disabled first, and then enabled upon data restoration.

Restoring Kafka Metadata

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, locate the desired success record and click **View** in the **Backup Path** column to display the task's backup path information and obtain the following details:

- **Backup Object**: indicates the backup data source.
- **Backup Path**: indicates the full path where the backup files are stored.

  Locate the correct path, and manually copy the full path of the backup files from the **Backup Path** column.

Step 4 On MRS Manager, choose **O&M** > **Backup and Restoration** > **Restoration Management**.

Step 5 Click **Create**.

Step 6 Set **Task Name** to the name of the restoration task.

Step 7 Select the cluster to be operated from **Recovery Object**.

Step 8 In the **Restoration Configuration** area, select **Kafka**.

Step 9 Select a backup directory type for **Path Type** of **Kafka**.

Table 9-72 Path for data restoration

Type	Parameter	Description
LocalDir	Source Path	Backup file to be restored. File name format: *Version_Data source_Task execution time*.tar.gz

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
LocalHDFS	Source Path	Full path of the HDFS directory storing backup files. Path format: *Backup path*/*Backup task name_Task creation time*/*Version_Data source_Task execution time*.tar.gz
	Source NameService Name	NameService name of the backup directory during restoration task execution. The default value is **hacluster**.
RemoteHDFS	Source NameService Name	NameService name of the backup data cluster. • You can set it to the NameService name (**haclusterX**, **haclusterX1**, **haclusterX2**, **haclusterX3**, or **haclusterX4**) of the built-in remote cluster. • You can also set it to the NameService name of a configured remote cluster.
	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Source NameNode IP Address	Service plane IP address of the active or standby NameNode in the standby cluster.
	Source Path	Full path of the HDFS directory storing backup data in the standby cluster. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.tar.gz
	Queue Name	Name of the YARN queue used for backup task execution. The name must be identical to that of a queue currently running properly in the cluster.
NFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.tar.gz

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
CIFS	IP Mode	IP version of the target IP address. The system automatically determines the IP version, such as IPv4 or IPv6, based on the cluster network type.
	Server IP Address	IP address of the NAS server.
	Port	Port number used by the CIFS protocol to connect to the NAS server. The default value is **445**.
	Username	Username configured during CIFS protocol setup.
	Password	Password configured during CIFS protocol setup.
	Source Path	Full path of the NAS server directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**
OBS (available in MRS 3.1.0 and later)	Source Path	Full path of the OBS directory storing backup files. Path format: *Backup path*/*Backup task name_Data source_Task creation time*/*Version_Data source_Task execution time*.**tar.gz**

Step 10 Click **OK**.

Step 11 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

- After the restoration is successful, the progress bar is in green.
- After the restoration is successful, the restoration task cannot be executed again.
- If the restoration task fails during the first execution, rectify the fault and click **Retry** to execute the task again.

  If the Kafka service is reinstalled and metadata is restored after data backup, or metadata is migrated to a new cluster, the Kafka Broker cannot be restarted. You can check the error details in the **/var/log/Bigdata/kafka/broker/server.log** file. An example error message is as follows:

  ```
  ERROR Fatal error during KafkaServer startup. Prepare to shutdown (kafka.server.KafkaServer)kafka.common.InconsistentClusterIdException: The Cluster ID kVSgfurUQFGGpHMTBqBPiw doesn't match stored clusterId Some(0Qftv9yBTAmf2iDPSIlk7g) in meta.properties. The broker is trying to join the wrong cluster. Configured zookeeper.connect may be wrong. at kafka.server.KafkaServer.startup(KafkaServer.scala:220) at kafka.server.KafkaServerStartable.startup(KafkaServerStartable.scala:44) at kafka.Kafka $.main(Kafka.scala:84) at kafka.Kafka.main(Kafka.scala)
  ```

  Check the value of **log.dirs** in the Kafka Broker configuration file **${BIGDATA_HOME}/FusionInsight_Current/*Broker/etc/server.properties**. The value is the Kafka data directory. Navigate to the Kafka data directory

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

and change the value **0Qftv9yBTAmf2iDPSlIk7g** of **cluster.id** in **meta.properties** to **kVSgfurUQFGGpHMTBqBPiw** (the latest value in the error log).

Perform the preceding modification on every node where a Kafka Broker is located. After completing the modification, restart the Kafka service.

----End

9.6.7 Managing MRS Cluster Backup and Restoration Tasks

This topic describes how to modify the parameters of a created backup task on Manager to meet changing service requirements. The parameters of restoration tasks can be viewed but not modified.

Impact on the System

After a backup task is modified, the new parameters take effect when the task is executed next time.

Prerequisites

- A backup task has been created.
- A new backup task policy has been planned based on the actual situation.

Modifying a Backup Task (MRS 3.x and Later)

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the task list, locate a specified task, click **Configure** in the **Operation** column to go to the configuration modification page.

On the displayed page, modify the following parameters:

- **Started**: time when the backup task is started for the first time.
- **Period**: interval between the next startup and the last execution of the backup task, supporting hourly or daily intervals.
- **Destination NameService Name**: NameService of the standby cluster.
- **Destination NameNode IP Address**: IP address of the standby cluster's NameNode service plane, supporting both active and standby nodes.
- **Destination Path**: HDFS directory for storing backup data of the standby cluster. The path cannot be configured as an HDFS hidden directory, such as a snapshot or recycle bin directory, nor as a default system directory, such as **/hbase** or **/user/hbase/backup**.
- **Maximum Number of Backup Copies**: number of backup file sets that can be retained in the backup directory.
- **Maximum Number of Recovery Points**: number of snapshots that can be retained in the cluster.
- **Maximum Number of Maps**: maximum number of maps for executing MapReduce tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Maximum Map Bandwidth (MB/s)**: maximum bandwidth of a single map.

After you change the value of **Target Path** for a backup task, full backup is performed by default when the task is executed for the first time.

Step 4 Click **OK** to save the settings.

----End

Modifying a Backup Task (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

Step 2 Choose **System** > **Back Up Data**.

Step 3 In the task list, locate a specified task, click **Modify** in the **Operation** column to go to the configuration modification page.

Step 4 Modify the following parameters on the displayed page:

- Manual backup:
  - **Target Path**: HDFS directory for storing the backup files.
  - **Maximum Number of Backup Copies**: number of backup file sets that can be retained in the backup directory.
- Periodic backup:
  - **Started**: time when the backup task is started for the first time.
  - **Period**: interval between the next startup and the last execution of the backup task, supporting hourly or daily intervals.
  - **Target Path**: HDFS directory for storing the backup files. The path cannot be an HDFS hidden directory, such as snapshot or recycle bin directory, or a default system directory.
  - **Maximum Number of Backup Copies**: number of backup file sets that can be retained in the backup directory.

When **Path Type** is set to **LocalHDFS**, **Target Path** is valid for modifying a backup task.

After you change the value of **Target Path** for a backup task, full backup is performed by default when the task is executed for the first time.

Step 5 Click **OK** to save the settings.

----End

Viewing Backup and Restoration Tasks (MRS 3.x and Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

Step 2 Choose **O&M** > **Backup and Restoration**.

Step 3 Click **Backup Management** or **Restoration Management**.

Step 4 In the task list, obtain the previous execution result in the **Task Status** and **Task Progress** column. Green indicates that the task is executed successfully, and red indicates that the execution fails.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 In the **Operation** column of a specified task in the task list, choose **More** > **View History** or click **View History** to view the historical record of backup and restoration task execution.

In the displayed window, click before a specified record to display log information about the execution.

You can also perform more maintenance operations listed in **Table 9-73** in the **Operation** column of a specified task.

Table 9-73 Maintenance and management operations

Operation Entry	Description
More > **Back Up Now** or **Start**	Start a backup or restoration task that is ready or fails to be executed. Executed restoration tasks cannot be repeatedly executed.
Configure	Modify parameters for a backup task.
Restore	After data backup, click **Restore** to quickly restore data.
More > **Back Up Now**	Execute a backup task immediately.
More > **Stop** or **Stop**	Stop a running backup or restoration task. After the task is stopped, its **Task Status** changes to **Stopped**.
More > **Delete** or **Delete**	Delete a backup or restoration task. After a task is deleted, the backup data is retained by default.
More > **Suspend**	Suspend a backup task. Only periodic backup tasks can be suspended. Suspended backup tasks are no longer executed automatically. If you suspend a backup task that is being executed, the task execution stops. To unlock the task and run it again, choose **More** > **Run Again**.
More > **Resume**	Enable automatic backup.
More > **View History** or **View History**	Switch to the task run log page to view the task running details and backup path.
View	Check the parameter settings of a restoration task.
Start	Run a restoration task.

----End

Viewing Backup and Restoration Tasks (MRS 2.x and Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

Step 2 Click **System**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click **Backup Management** or **Restoration Management**.

Step 4 In the task list, obtain the previous execution result in the **Task Progress** column. Green indicates that the task is executed successfully, and red indicates that the execution fails.

Step 5 In the row containing the specified backup task, choose **More** > **View History** in the **Operation** column to display the task's historical execution records.

In the displayed window, click **View** in the **Details** column. The task execution logs and paths are displayed.

NOTE

Other related operations are as follows:

- Viewing a restoration task

  In the **Operation** column of the specified task in the task list, click **View Details** to view the restoration task. You can only view but cannot modify the parameters of a restoration task.

- Executing a backup or restoration task

  In the task list, locate a specified task and click **Start** in the **Operation** column to start a backup or restoration task that is ready or fails to be executed. Executed restoration tasks cannot be repeatedly executed.

- Stopping a backup task

  In the task list, locate a specified task and choose **More** > **Stop** in the **Operation** column to stop a backup task that is running.

- Deleting a backup or restoration task

  In the **Operation** column of the specified task in the task list, choose **More** > **Delete** to delete the backup or restoration task. After a task is deleted, the backup data is retained by default.

- Suspending a backup task

  In the task list, locate a specified task and choose **More** > **Suspend** in the **Operation** column to suspend the backup task. Only periodic backup tasks can be suspended. Suspended backup tasks are no longer executed automatically. When you suspend a backup task that is being executed, the task execution stops. To cancel the suspension status of a task, choose **More** > **Resume**.

----End

9.6.8 Using HDFS Snapshots to Quickly Restore Component Service Data

Scenarios

When DistCp is used to back up data, the backup snapshot is saved to HDFS of the active cluster. FusionInsight Manager supports using the local snapshot for quick data restoration, requiring less time than restoring data from the standby cluster.

Use FusionInsight Manager and the snapshots on HDFS of the active cluster to create a local quick restoration task and execute the task.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Restoring Component Service Data

Step 1 Log in to MRS Manager.

Step 2 Choose **O&M** > **Backup and Restoration** > **Backup Management**.

Step 3 In the backup task list, locate a created task and click **Restore** in the **Operation** column.

Step 4 Check whether the system displays "No data is available for quick restoration. Create a task on the restoration management page to restore data."

- If yes, click **OK** to close the dialog box. No backup data snapshot is created in the active cluster, and no further action is required.
- If no, go to **Step 5** to create a local quick restoration task.

Metadata does not support quick restoration.

Step 5 Set **Name** to the name of the local quick restoration task.

Step 6 Set **Configuration** to a data source.

Step 7 Set **Recovery Point List** to a recovery point that contains the backup data.

Step 8 Set **Queue Name** to the name of the Yarn queue used in the task execution. The name must be identical to that of a queue currently running properly in the cluster.

Step 9 Set **Data Configuration** to the object to be recovered.

Step 10 Click **Verify**, and wait for the system to display "The restoration task configuration is verified successfully."

Step 11 Click **OK**.

Step 12 In the restoration task list, locate the row containing the created task, and click **Start** in the **Operation** column to execute the restoration task.

After the task is complete, **Task Status** of the task is displayed as **Successful**.

----End

9.7 MRS Cluster Patching

9.7.1 Viewing Patch Information for an MRS Cluster

View patch information for cluster components. If a cluster component, such as Hadoop or Spark, is abnormal, download the patch version to rectify the fault.

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to enter the cluster details page.

Step 3 Click **Patches**. On the displayed page, view patch information for the MRS cluster.

- Patch Name: name of the patch package
- Published: time when the patch package is released

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Status: patch status
- Patch Description: patch version description
- Operation: patch installation or uninstallation

----End

9.7.2 Patching an MRS Cluster

Install the cluster version patch as required if you obtain patch information from:

- The message center service
- On the details page of an MRS cluster on the MRS console, choose **Patches** > **Cluster Component Patches** to view the patches that can be installed.

Preparing for Patch Installation

- For details about how to check the cluster status, see **Performing a Health Check for an MRS Cluster**. Exceptions such as cluster node faults and hard disk faults may cause patch installation and uninstallation failures. Before you install or uninstall the patch, ensure that the cluster is healthy.
- Click **Patches** and then **Cluster Component Patches**, check the **Patch Description** column of the target patch, and read the patch description carefully to understand the patch installation procedure and impact.
- MRS 2.x and earlier versions, MRS 3.1.5 and later versions, and MRS 3.2.0-LTS and later versions support online patch installation.

Installing a Patch

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to enter the cluster details page.

Step 3 In the **Patches** tab, click **Cluster Component Patches**. In the operation list, click **Install** next to the patch you want to install.

Step 4 In the displayed dialog box, select **I have read Patch Description and understood that this operation may restart services.** Click **Yes** and wait until the patch is successfully installed.

Step 5 Check the patch status. Restart the components and install the client patch according to the patch description.

- If a host in the cluster is isolated, the patch will not be installed on that host. After installation completes, the patch status will be reported as partially installed. After an isolated node is restored and isolation is canceled, you can install the patch again. In this case, the patch is applied only to the node where isolation has been canceled. For versions earlier than MRS 3.x, perform operations by referring to **Patching Hosts Isolated in an MRS Cluster**.
- To install client patches in batches for MRS 3.3.0 or later clusters, perform the following operations:

  a. Log in to FusionInsight Manager of the MRS cluster.

     For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

b. Choose **Cluster** > **Client Management**.

c. On the **Client Management** page, select the clients where the patches you want to install and choose **More** > **Batch Install Patches**.

d. In the displayed dialog box, select the checkbox to confirm the operation impact and click **OK** to install the MRS patches in batches.

During patch installation, clients cannot run properly and services on the clients may be interrupted.

----End

Uninstalling a Patch

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to enter the cluster details page.

Step 3 In the **Patches** tab, click **Cluster Component Patches**. In the operation list, click **Uninstall** next to the patch you want to uninstall.

Step 4 In the displayed dialog box, select the confirmation check box and click **Yes**. Wait until the patch is successfully uninstalled.

Step 5 Restart the component and uninstall the client patch according to the patch description.

If a host in the cluster is isolated, the patch will not be uninstalled on that host. After uninstallation completes, the patch status will be reported as partially uninstalled. After the isolated node is restored and isolation is canceled, you can uninstall the patch again. In this case, the patch is removed only from the node whose isolation has been canceled. For versions earlier than MRS 3.x, perform operations by referring to **Patching Hosts Isolated in an MRS Cluster**.

----End

9.7.3 Applying Rolling Patches for an MRS Cluster

The rolling patch function indicates that patches are installed or uninstalled for one or more services in a cluster by performing a rolling service restart (restarting services or instances in batches), without interrupting the services or within a minimized service interruption interval. Services in a cluster are divided into the following three types based on whether they support rolling patch:

- Services supporting rolling patch installation or uninstallation: All businesses or part of them (varying depending on different services) of the services are not interrupted during patch installation or uninstallation.
- Services not supporting rolling patch installation or uninstallation: Businesses of the services are interrupted during patch installation or uninstallation.
- Services with some roles supporting rolling patch installation or uninstallation: Some businesses of the services are not interrupted during patch installation or uninstallation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Notes and Constraints

In **MRS 3.x**, you cannot perform operations in this section on the management console.

Table 9-74 provides services and instances that support or do not support rolling restart in the MRS cluster.

Table 9-74 Support for rolling restarts on services and instances

Service	Instance	Support for Rolling Restart
Alluxio	AlluxioJobMaster	Yes
	AlluxioMaster	
Flink	FlinkResource	No
	FlinkServer	
Flume	Flume	Yes
	MonitorServer	
HBase	HMaster	Yes
	RegionServer	
	ThriftServer	
	RESTServer	
HDFS	NameNode	Yes
	Zkfc	
	JournalNode	
	HttpFS	
	DataNode	
Hive	MetaStore	Yes
	WebHCat	
	HiveServer	
Hue	Hue	No
Impala	Impalad	No
	StateStore	
	Catalog	
Kafka	Broker	Yes
	KafkaUI	No

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Instance	Support for Rolling Restart
Kudu	KuduTserver	Yes
	KuduMaster	
Loader	Sqoop	No
Mapreduce	JobHistoryServer	Yes
Oozie	oozie	No
Presto	Coordinator	Yes
	Worker	
Spark	JobHistory	Yes
	JDBCServer	
	SparkResource	
Storm	Nimbus	Yes
	UI	
	Supervisor	
	Logviewer	
Tez	TezUI	No
Yarn	ResourceManager	Yes
	NodeManager	
Zookeeper	Quorumpeer	Yes

Installing a Rolling Patch

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to enter the cluster details page.

Step 3 On the **Patches** page, click **Install** in the **Operation** column.

Step 4 On the **Warning** page, enable or disable **Rolling Patch**.

- Enabling the rolling patch installation function: Services are not stopped before patch installation, and rolling service restart is performed after the patch installation. This process minimizes the impact on cluster services but requires more time than common patch installation.
- Disabling the rolling patch installation function: All services are stopped before patch installation and then restarted after the patch installation. This process temporarily interrupts the cluster and its services but requires less time than rolling patch installation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The rolling patch installation function is unavailable in clusters with fewer than two Master nodes or fewer than three Core nodes.

Step 5 Click **Yes** to install the target patch.

Step 6 View the patch installation progress.

1. Access MRS Manager. For details, see **Accessing MRS Manager**.
2. Choose **System** > **Manage Patch**. On the **Manage Patch** page, you can view the patch installation progress.

NOTE

For the isolated host nodes in the cluster, follow instructions in **Patching Hosts Isolated in an MRS Cluster** to restore the patch.

----End

Uninstalling a Rolling Patch

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page displayed by default, click the name of the target cluster to enter the cluster details page.

Step 3 On the **Patches** page, click **Uninstall** in the **Operation** column.

Step 4 On the **Warning** page, enable or disable **Rolling Patch**.

- Enabling the rolling patch uninstallation function: Services are not stopped before patch uninstallation, and rolling service restart is performed after the patch uninstallation. This process minimizes the impact on cluster services but requires more time than common patch uninstallation.
- Disabling the rolling patch uninstallation function: All services are stopped before patch uninstallation and then restarted after the patch uninstallation. This process temporarily interrupts the cluster and its services but requires less time than rolling patch uninstallation.
- Only patches installed in rolling mode can be uninstalled in rolling mode.

Step 5 Click **Yes** to uninstall the target patch.

Step 6 View the patch uninstallation progress.

1. Access MRS Manager. For details, see **Accessing MRS Manager**.
2. Choose **System** > **Manage Patch**. On the **Manage Patch** page, you can view the patch uninstallation progress.

NOTE

For the isolated host nodes in the cluster, follow instructions in **Patching Hosts Isolated in an MRS Cluster** to restore the patch.

----End

9.7.4 Patching Hosts Isolated in an MRS Cluster

If some hosts are isolated in a cluster, perform the following operations to restore patches for these isolated hosts after patch installation on other hosts in the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cluster. After patch restoration, versions of the isolated host nodes are consistent with those are not isolated.

Operations in this section cannot be performed on the management console of MRS 3.x. This section applies only to versions earlier than 3.x.

Restoring Patches on Isolated Hosts

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System > Manage Patch**. The **Manage Patch** page is displayed.

Step 3 In the **Operation** column, click **View Details**.

Step 4 On the patch details page, select host nodes whose **Status** is **Isolated**.

Step 5 Click **Select and Restore** to restore the isolated host nodes.

----End

9.8 Viewing Logs of an MRS Cluster

9.8.1 Overview of MRS Cluster Logs

Log Description

MRS cluster logs are stored in the **/var/log/Bigdata** directory. The following table lists the log types.

Table 9-75 Log types

Type	Description
Installation log	Installation logs record information about FusionInsight Manager, cluster, and service installation to help users locate installation errors.
Run logs	Run logs record the debugging information, status changes, potential problems, and error information generated during the running of services.
Audit logs	Audit logs record information about users' activities and operation instructions, which can be used to locate fault causes in security events and determine who are responsible for these faults.

The following table lists the MRS log directories.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-76 Log directories

File Directory	Log Content
/var/log/Bigdata/audit	Component audit log.
/var/log/Bigdata/controller	Log collecting script log. Controller process log. Controller monitoring log.
/var/log/Bigdata/dbservice	DBService log.
/var/log/Bigdata/flume	Flume log.
/var/log/Bigdata/hbase	HBase log.
/var/log/Bigdata/hdfs	HDFS log.
/var/log/Bigdata/hive	Hive log.
/var/log/Bigdata/hetuengine	HetuEngine logs.
/var/log/Bigdata/httpd	HTTPD log.
/var/log/Bigdata/hue	Hue log.
/var/log/Bigdata/kerberos	Kerberos log.
/var/log/Bigdata/ldapclient	LDAP client log.
/var/log/Bigdata/ldapserver	LDAP server log.
/var/log/Bigdata/loader	Loader log.
/var/log/Bigdata/logman	logman script log management log.
/var/log/Bigdata/mapreduce	MapReduce log.
/var/log/Bigdata/nodeagent	NodeAgent log.
/var/log/Bigdata/okerberos	OMS Kerberos log.
/var/log/Bigdata/oldapserver	OMS LDAP log.
/var/log/Bigdata/metric_agent	MetricAgent run log.
/var/log/Bigdata/omm	**oms**: complex event processing log, alarm service log, HA log, authentication and authorization management log, and monitoring service run log of the omm server. **oma**: installation log and run log of the omm agent. **core**: dump log generated when the omm agent and the HA process are suspended.
/var/log/Bigdata/spark2x	Spark2x log.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

File Directory	Log Content
/var/log/Bigdata/sudo	Log generated when the **sudo** command is executed by user **omm**.
/var/log/Bigdata/timestamp	Time synchronization management log.
/var/log/Bigdata/tomcat	Tomcat log.
/var/log/Bigdata/watchdog	Watchdog log.
/var/log/Bigdata/yarn	YARN log.
/var/log/Bigdata/zookeeper	ZooKeeper log.
/var/log/Bigdata/oozie	Oozie log.
/var/log/Bigdata/kafka	Kafka log.
/var/log/Bigdata/storm	Storm log.
/var/log/Bigdata/iotdb	IoTDB log.
/var/log/Bigdata/cdl	CDL log.
/var/log/Bigdata/upgrade	OMS upgrade log.
/var/log/Bigdata/update-service	Upgrade service log.
/var/log/Bigdata/patch	Patch log.

Run logs

Table 9-77 describes the running information recorded in run logs.

Table 9-77 Running information

Run Log	Description
Installation preparation log	Records information about preparations for the installation, such as the detection, configuration, and feedback operation information.
Process startup log	Records information about the commands executed during the process startup.
Process startup exception log	Records information about exceptions during process startup, such as dependent service errors and insufficient resources.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Run Log	Description
Process run log	Records information about the process running information and debugging information, such as function entries and exits as well as cross-module interface messages.
Process running exception log	Records errors that cause process running errors, for example, the empty input objects or encoding or decoding failure.
Process running environment log	Records information about the process running environment, such as resource status and environment variables.
Script logs	Records information about the script execution process.
Resource reclamation log	Records information about the resource reclaiming process.
Uninstallation clearing logs	Records information about operations performed during service uninstallation, such as directory deletion and execution time

Audit logs

Audit information recorded in audit logs includes FusionInsight Manager audit information and component audit information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-78 Audit information of FusionInsight Manager

Operation Type	Operation
User management	Creating a user Modifying a user Deleting a user Creating a user group Modifying a user group Deleting a user group Adding a role Modifying a role Deleting a role Changing a password policy Changing a password Resetting a password User login User logout Unlocking the screen Downloading the authentication credential Unauthorized operation Unlocking a user account Locking a user account Locking the screen Exporting user information Exporting a user group Exporting a role

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
Cluster	Starting a cluster Stopping a cluster Restarting a cluster Performing a rolling restart of a cluster Restarting all expired instances Saving configurations Synchronizing cluster configurations Customizing cluster monitoring indicators Configuring monitoring dumping Saving monitoring thresholds Downloading a client configuration file Configuring the northbound Syslog interface Configuring the northbound SNMP API Clearing alarm through SNMP Adding a trap target through SNMP Deleting a trap target through SNMP Checking alarms through SNMP Synchronizing alarms through SNMP Creating a threshold template Deleting a threshold template Applying a threshold template Saving cluster monitoring configuration data Exporting configuration data Importing cluster configuration data Exporting an installation template Modifying a threshold template Canceling the application of a threshold template Masking alarms Sending an alarm Changing the OMS database password Resetting the component database password Restarting OMM and Controller Starting the health check of a cluster Importing a certificate file Configuring SSO information Deleting historical health check reports Modifying cluster attributes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
	Running maintenance commands in synchronous mode Running maintenance commands in asynchronous mode Customizing report monitoring indicators Exporting report monitoring data Running a command using SNMP in asynchronous mode Restarting the Web service Customizing monitoring indicators for static resource pools Exporting monitoring data of a static resource pool Customizing dashboard monitoring metrics Stopping a task Restoring configurations Modifying domain and mutual trust configurations Modifying system parameters Making a cluster enter the maintenance mode Making a cluster exit the maintenance mode Making OMS enter the maintenance mode Making OMS exit the maintenance mode Exiting the maintenance mode in batches Modifying OMS configurations Enabling threshold alarms Synchronizing all cluster configurations

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
Service	Starting a service Stopping a service Synchronizing service configurations Refreshing a service queue Customizing service monitoring indicators Restarting a service Performing a rolling restart of a service Exporting service monitoring data Importing service configuration data Starting the health check of a service Configuring the service Uploading a configuration file Downloading a configuration file Synchronizing instance configurations Commissioning an instance Decommissioning an instance Starting an instance Stopping an instance Customizing instance monitoring indicators Restarting an instance Performing a rolling restart of an instance Exporting instance monitoring data Importing instance configuration data Creating an instance group Modifying an instance group Deleting an instance group Moving an instance to another instance group Making a service enter the maintenance mode Making a service exit the maintenance mode Changing the display name of a service Modifying service association Downloading monitoring data Masking alarms Unmasking alarms Exporting report data of a service Adding custom parameters for a report Modifying custom parameters of a report Deleting custom parameters of a report

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
	Switching over controller nodes Adding a mount table Modifying a mount table
Host	Setting a node rack Starting all roles Stopping all roles Isolating a host Canceling host isolation Customizing host monitoring indicators Exporting host monitoring data Making a host enter the maintenance mode Making a host exit the maintenance mode Exporting basic host information Exporting host distribution report data Exporting host trend report data Exporting host cluster report data Exporting service report data Customizing host cluster monitoring metrics Customizing host trend monitoring metrics
Alarm	Exporting alarms Clearing alarms Exporting events Clearing alarms in batches
Log collection	Collecting log files Downloading log files Collecting service stack information Collecting instance stack information Preparing service stack information Preparing instance stack information Clearing service stack information Clearing instance stack information
Audit logs	Modifying audit dump configurations Exporting audit logs

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
Backup and restoration	Creating a backup task Executing a backup task Executing backup tasks in batches Stopping a backup task Deleting a backup task Modifying a backup task Locking a backup task Unlocking a backup task Creating a restoration task Executing a backup restoration task Stopping a restoration task Retrying a restoration task Deleting a restoration task

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
Multi-tenant management	Saving the static configuration Creating a tenant Deleting a tenant Associating a service with a tenant Deleting a service from a tenant Configuring resources Creating resources Deleting resources Adding a resource pool Modifying a resource pool Deleting a resource pool Restoring tenant data Modifying a tenant's global configurations Modifying queue configurations for a capacity scheduler Modifying queue configurations for a super scheduler Modifying resource distribution for a capacity scheduler Clearing resource distribution for a capacity scheduler Modifying resource distribution for a super scheduler Clearing resource distribution for a super scheduler Adding a resource catalog Modifying a resource catalog Deleting a resource catalog Customizing tenant monitoring metrics

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Operation Type	Operation
Health check	Starting health check for a cluster Starting health check for a service Starting health check for a host Starting health check for OMS Starting health check for a system Updating health check configurations Exporting health check reports Exporting cluster health check results Exporting service health check results Exporting host health check results Deleting historical health check reports Exporting historical health check reports Downloading health check reports

Table 9-79 Component audit information

Audit Log	Operation Type	Operation
CDL audit log	Service operation	Creating a link Deleting a link Creating a job Starting a job Deleting a job
IoTDB audit log	Maintenance management	Granting permissions Revoking permissions Recording authentication and login information
	Service operation	Deleting a time series, partition, function, or index Modifying a time series
ClickHouse audit log	Maintenance management	Granting permissions Revoking permissions Recording authentication and login information

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Audit Log	Operation Type	Operation
	Service operation	Creating a database or table Inserting, deleting, querying, and migrating data
DBService audit log	Maintenance management	Performing backup restoration operations
HBase audit log	Data definition language (DDL) statement	Creating a table Deleting a table Modifying a table Adding a column family Modifying a column family Deleting a column family Enabling a table Disabling a table Modify the user information Changing a password User login
	Data manipulation language (DML) statement	Putting data (to the **hbase:meta**, **_ctmeta_**, and **hbase:acl** tables) Deleting data (from the **hbase:meta**, **_ctmeta_**, and **hbase:acl** tables) Checking and putting data (to the **hbase:meta**, **_ctmeta_**, and **hbase:acl** tables) Checking and deleting data (from the **hbase:meta**, **_ctmeta_**, and **hbase:acl** tables)
	Permission control	Assigning permissions to a user Canceling permission assigning

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Audit Log	Operation Type	Operation
HDFS audit log	Permissions management	Managing permissions on files or folders Managing permissions on owner information files or folders
	File operation	Creating a folder Creating a file Opening a file Appending file content Changing a file name Deleting a file or folder Setting time property of a file Setting the number of file copies Merging files Checking the file system File links
Hive audit logs	Metadata operation	Defining metadata, such as creating databases and tables Deleting metadata, such as deleting databases and tables Modifying metadata, such as adding columns and renaming tables Importing and exporting metadata
	Data maintenance	Loading data to a table Inserting data into a table
	Permissions management	Creating or deleting roles Granting/Reclaiming roles Granting/Reclaiming permissions
Hue audit log	Service startup	Starting Hue
	User operation	User login User logout

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Audit Log	Operation Type	Operation
	Task operation	Creating a job Modifying a job Deleting a job Submitting a task Saving a task Updating the status of a task
KrbServer audit log	Maintenance management	Changing the password of a Kerberos account Adding a Kerberos account Deleting a Kerberos account Authenticating a user
LdapServer audit log	Maintenance management	Adding an operating system user Adding a user group Adding a user to user group Deleting a user Deleting a group
Loader audit log	Security management	User login
	Metadata management	Querying connector information Querying a framework Querying step information
	Managing data source connections	Querying a data source connection Adding a data source connection Updating a data source connection Deleting a data source connection Activating a data source connection Disabling a data source connection

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Audit Log	Operation Type	Operation
	Job management	Querying a job Creating a Job Updating a Job Deleting a job Activating a job Disabling a job Querying all execution records of a job Querying the latest execution record of a job Submitting a job Stopping a job
MapReduce audit log	Application running	Starting a Container request Stopping a Container request After Container request is completed, the status of the request is displayed as succeeded. After Container request is completed, the status of the request is displayed as failed. After Container request is completed, the status of the request is displayed as suspended. Submitting a task Ending a task
Oozie audit log	Task management	Submitting a task Starting a task Killing a task Suspending a task Resuming a task Re-running a task

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Audit Log	Operation Type	Operation
Spark audit log	Metadata operation	Defining metadata, such as creating databases and tables Deleting metadata, such as deleting databases and tables Modifying metadata, such as adding columns and renaming tables Importing and exporting metadata
	Data maintenance	Loading data to a table Inserting data into a table
Storm audit log	Nimbus	Submitting a topology Stopping a topology Reallocating a topology Deactivating a topology Activating a topology
	UI	Stopping a topology Reallocating a topology Deactivating a topology Activating a topology
YARN audit log	Job submission	Submitting a job to a queue
ZooKeeper audit log	Permissions management	Setting the access permission to Znode
	Znode operation	Creating a Znode Deleting a Znode Configuring Znode data

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Audit Log	Operation Type	Operation
HetuEngine audit log	Job management	Adding an external data source Deleting an external data source Modifying an external data source Creating a compute instance Starting a compute instance Stopping a compute instance Deleting a compute instance Querying a compute instance Modifying compute instance configurations

MRS audit logs are stored in the database. You can view and export audit logs on the **Audit** page.

The following table lists the directories to store component audit logs. Audit log files of some components are stored in **/var/log/Bigdata/audit**, such as HDFS, HBase, MapReduce, Hive, Hue, YARN, Storm, and ZooKeeper. The component audit logs are automatically compressed and backed up to **/var/log/Bigdata/audit/bk** at 03:00 every day. A maximum of latest 90 compressed backup files are retained, and the backup time cannot be changed.

Audit log files of other components are stored in the component log directory.

Table 9-80 Directory for storing component audit logs

Component	Audit Log Directory
DBService	/var/log/Bigdata/audit/dbservice/dbservice_audit.log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Audit Log Directory
HBase	/var/log/Bigdata/audit/hbase/hm/hbase-audit-hmaster.log /var/log/Bigdata/audit/hbase/hm/hbase-ranger-audit-hmaster.log /var/log/Bigdata/audit/hbase/rs/hbase-audit-regionserver.log /var/log/Bigdata/audit/hbase/rs/hbase-ranger-audit-regionserver.log /var/log/Bigdata/audit/hbase/rt/hbase-audit-restserver.log /var/log/Bigdata/audit/hbase/ts/hbase-audit-thriftserver.log
HDFS	/var/log/Bigdata/audit/hdfs/nn/hdfs-audit-namenode.log /var/log/Bigdata/audit/hdfs/nn/ranger-plugin-audit.log /var/log/Bigdata/audit/hdfs/dn/hdfs-audit-datanode.log /var/log/Bigdata/audit/hdfs/jn/hdfs-audit-journalnode.log /var/log/Bigdata/audit/hdfs/zkfc/hdfs-audit-zkfc.log /var/log/Bigdata/audit/hdfs/httpfs/hdfs-audit-httpfs.log /var/log/Bigdata/audit/hdfs/router/hdfs-audit-router.log
HetuEngine	/var/log/Bigdata/audit/hetuengine/hsbroker/hsbroker-audit.log.0 /var/log/Bigdata/audit/hetuengine/hsconsole/hsconsole-audit.log.0 /var/log/Bigdata/audit/hetuengine/hsfabric/hsfabric-audit.log.0 hdfs://hacluster/hetuserverhistory/*Tenant name*/coordinator/application_*ID*/container_*ID*/*yyyyMMdd*/hetuserver-engine-audit.log hdfs://hacluster/hetuserverhistory/*Tenant name*/coordinator or worker/application_*ID*/container_*ID*/*yyyyMMdd*/server.log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Audit Log Directory
Hive	/var/log/Bigdata/audit/hive/hiveserver/hive-audit.log /var/log/Bigdata/audit/hive/hiveserver/hive-rangeraudit.log /var/log/Bigdata/audit/hive/metastore/metastore-audit.log /var/log/Bigdata/audit/hive/webhcat/webhcat-audit.log
Hue	/var/log/Bigdata/audit/hue/hue-audits.log
Kafka	/var/log/Bigdata/audit/kafka/audit.log
Loader	/var/log/Bigdata/loader/audit/default.audit
CDL	/var/log/Bigdata/audit/cdl/service/cdl-audit.log
MapReduce	/var/log/Bigdata/audit/mapreduce/jobhistory/mapred-audit-jobhistory.log
Oozie	/var/log/Bigdata/audit/oozie/oozie-audit.log
Spark2x	/var/log/Bigdata/audit/spark2x/jdbcserver/jdbcserver-audit.log /var/log/Bigdata/audit/spark2x/jdbcserver/ranger-audit.log /var/log/Bigdata/audit/spark2x/jobhistory/jobhistory-audit.log
Storm	/var/log/Bigdata/audit/storm/logviewer/audit.log /var/log/Bigdata/audit/storm/nimbus/audit.log /var/log/Bigdata/audit/storm/supervisor/audit.log /var/log/Bigdata/audit/storm/ui/audit.log
Yarn	/var/log/Bigdata/audit/yarn/rm/yarn-audit-resourcemanager.log /var/log/Bigdata/audit/yarn/rm/ranger-plugin-audit.log /var/log/Bigdata/audit/yarn/nm/yarn-audit-nodemanager.log
ZooKeeper	/var/log/Bigdata/audit/zookeeper/quorumpeer/zk-audit-quorumpeer.log
IoTDB	/var/log/Bigdata/audit/iotdb/iotdbserver/log_audit.log

Manager Logs

Log path: Manager log files are stored in **/var/log/Bigdata/***Manager* by default.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- ControllerService: **/var/log/Bigdata/controller/** (OMS installation and run logs)
- Httpd: **/var/log/Bigdata/httpd** (httpd installation and run logs)
- logman: **/var/log/Bigdata/logman** (log packaging tool logs)
- NodeAgent: **/var/log/Bigdata/nodeagent** (NodeAgent installation and run logs)
- okerberos: **/var/log/Bigdata/okerberos** (okerberos installation and run logs)
- oldapserver: **/var/log/Bigdata/oldapserver** (oldapserver installation and run logs)
- MetricAgent: **/var/log/Bigdata/metric_agent** (MetricAgent run logs)
- omm: **/var/log/Bigdata/omm** (omm installation and run logs)
- timestamp: **/var/log/Bigdata/timestamp** (NodeAgent startup time logs)
- tomcat: **/var/log/Bigdata/tomcat** (Web process logs)
- Watchdog: **/var/log/Bigdata/watchdog** (watchdog logs)
- Upgrade: **/var/log/Bigdata/upgrade** (OMS upgrade logs)
- UpdateService: **/var/log/Bigdata/update-service** (upgrade service logs)
- Sudo: **/var/log/Bigdata/sudo** (sudo script execution logs)
- OS: **/var/log/***message file* (OS system logs)
- OS Performance: **/var/log/osperf** (OS performance statistics logs)
- OS Statistics: **/var/log/osinfo/statistics** (OS parameter configuration logs)

Log archiving rules:

Manager logs are automatically compressed and archived. When a log file exceeds 10 MB in size, the file is automatically compressed by default. The file is named as follows after compression: <*Original log name*>-<*yyyy-mm-dd_hh-mm-ss*>.[*ID*].**log.zip**. A maximum of 20 latest compressed files are reserved.

Table 9-81 Manager logs

Type	Log File Name	Description
Controller run log	controller.log	Log file that records component installation, upgrade, configuration, monitoring, alarm reporting, and routine O&M operations
	controller_client.log	Run log file of REST APIs
	acs.log	ACS run log file
	acs_spnego.log	spnego user log in ACS
	aos.log	AOS run log
	plugin.log	AOS plug-in log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	backupplugin.log	Log file that records backup and restoration operations
	controller_config.log	Configuration run log
	controller_nodesetup.log	Controller loading task log
	controller_root.log	System log of the Controller process
	controller_trace.log	Log that records the remote procedure call (RPC) communication between Controller and NodeAgent
	controller_monitor.log	Monitoring log
	controller_fsm.log	State machine log
	controller_alarm.log	Controller alarm log
	controller_backup.log	Controller backup and restoration log
	install.log, restore_package.log, installPack.log, distributeAdapterFiles.log, and install_os_optimization.log	OMS installation log
	oms_ctl.log	OMS startup and stop log
	preInstall_client.log	Preprocessing log file before client installation
	installntp.log	NTP installation log
	modify_manager_param.log	Manager parameter modification log
	backup.log	OMS backup script run log
	supressionAlarm.log	Alarm script run log
	om.log	OM certificate generation log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	backupplugin_ctl.log	Startup log of the backup and restoration plug-in process
	getLogs.log	Run log of the collection log script
	backupAuditLogs.log	Run log of the audit log backup script
	certStatus.log	Log that records regular certificate checks
	distribute.log	Certificate distribution log
	ficertgenetrate.log	Certificate replacement logs, including logs of level-2 certificates, CAS certificates, and httpd certificates
	genPwFile.log	Log that records the generation of certificate password files
	modifyproxyconf.log	Log that records the modification of the HTTPD proxy configuration
	importTar.log	Log that records the process of importing certificates into the trust library
Httpd	install.log	Httpd installation log
	access_log, error_log	Httpd run log
logman	logman.log	Log packaging tool log
NodeAgent	install.log, install_os_optimization.log	NodeAgent installation log
	installntp.log	NTP installation log
	start_ntp.log	NTP startup log
	ntpChecker.log	NTP check log
	ntpMonitor.log	NTP monitoring log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	heartbeat_trace.log	Log that records heartbeats between NodeAgent and Controller
	alarm.log	Alarm log
	monitor.log	Monitoring log
	nodeagent_ctl.log, start-agent.log	NodeAgent startup log
	agent.log	NodeAgent run log
	cert.log	Certificate log
	agentplugin.log	Agent plug-in running status monitoring log
	omaplugin.log	OMA plug-in run log
	diskhealth.log	Disk health check log
	supressionAlarm.log	Alarm script run log
	updateHostFile.log	Host list update log
	collectLog.log	Run log of the node log collection script
	host_metric_collect.log	Host index collection run log
	checkfileconfig.log	Run log file of file permission check
	entropycheck.log	Entropy check run log
	timer.log	Log of periodic node scheduling
	pluginmonitor.log	Component monitoring plug-in log
	agent_alarm_py.log	Log that records alarms upon insufficient NodeAgent file permission
okerberos	addRealm.log and modifyKerberosRealm.log	Realm handover log
	checkservice_detail.log	Okerberos health check log
	genKeytab.log	keytab generation log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	KerberosAdmin_genConfigDetail.log	Run log file of **kadmin.conf** generated during start of the kadmin process
	KerberosServer_genConfigDetail.log	Run log file of **krb5kdc.conf** generated during start of the krb5kdc process
	oms-kadmind.log	Run log of the kadmin process
	oms_kerberos_install.log and postinstall_detail.log	Okerberos installation log
	oms-krb5kdc.log	Run log of the krbkdc process
	start_detail.log	Okerberos startup log
	realmDataConfigProcess.log	Log file that records the rollback upon a realm handover failure
	stop_detail.log	Okerberos stop log
oldapserver	ldapserver_backup.log	Oldapserver backup log
	ldapserver_chk_service.log	Oldapserver health check log
	ldapserver_install.log	Oldapserver installation log
	ldapserver_start.log	Oldapserver startup log
	ldapserver_status.log	Log that records the status of the Oldapserver process
	ldapserver_stop.log	Oldapserver stop log
	ldapserver_wrap.log	Oldapserver service management log
	ldapserver_uninstall.log	Oldapserver uninstallation log
	restart_service.log	Oldapserver restart log
	ldapserver_unlockUser.log	Log file that records information about unlocking LDAP users and managing accounts

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
metric_agent	gc.log	MetricAgent JVM GC log file
	metric_agent.log	Run log file of MetricAgent
	metric_agent_qps.log	Log file that records MetricAgent Internal queue length and QPS information
	metric_agent_root.log	All run log files of MetricAgent
	start.log	Log file that records information about the MetricAgent startup and stop
omm	omsconfig.log	OMS configuration log
	check_oms_heartbeat.log	OMS heartbeat log
	monitor.log	OMS monitoring log
	ha_monitor.log	HA_Monitor operation log
	ha.log	HA operation log
	fms.log	Alarm log
	fms_ha.log	HA alarm monitoring log
	fms_script.log	Alarm control log
	config.log	Alarm configuration log
	iam.log	IAM log
	iam_script.log	IAM control log
	iam_ha.log	IAM HA monitoring log
	config.log	IAM configuration log
	operatelog.log	IAM operation log
	heartbeatcheck_ha.log	OMS heartbeat HA monitoring log
	install_oms.log	OMS installation log
	pms_ha.log	HA monitoring log
	pms_script.log	Monitoring control log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	config.log	Monitoring configuration log
	plugin.log	Monitoring plug-in run log
	pms.log	Monitoring log
	ha.log	HA run log
	cep_ha.log	CEP HA monitoring log
	cep_script.log	CEP control log
	cep.log	CEP log
	config.log	CEP configuration log
	omm_gaussdba.log	GaussDB HA monitoring log
	gaussdb-<SERIAL>.log	GaussDB run log
	gs_ctl-<DATE>.log	GaussDB control log archive log
	gs_ctl-current.log	GaussDB control log
	gs_guc-current.log	GaussDB operation log
	encrypt.log	Omm encryption log
	omm_agent_ctl.log	OMA control log
	oma_monitor.log	OMA monitoring log
	install_oma.log	OMA installation log
	config_oma.log	OMA configuration log
	omm_agent.log	OMA run log
	acs.log	ACS resource log
	aos.log	AOS resource log
	controller.log	Controller resource log
	floatip.log	Floating IP address resource log
	ha_ntp.log	NTP resource log
	httpd.log	Httpd resource log
	okerberos.log	Okerberos resource log
	oldap.log	OLdap resource log

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	tomcat.log	Tomcat resource log
	send_alarm.log	Run log of the HA alarm sending script of the management node
	feed_watchdog.log	feed_watchdog resource log
timestamp	restart_stamp	NodeAgent start time log
tomcat	cas.log and localhost_access_cas_log.log	CAS run log
	catalina.log, catalina.out, host-manager.log, localhost.log, and manager.log	Tomcat run log
	localhost_access_web_log.log	Log that records the access to REST APIs of FusionInsight Manager
	web.log	Run log of the web process
	northbound_ftp_sftp.log and snmp.log	Northbound log
	perfStats.log	Performance statistics log file
watchdog	watchdog.log and feed_watchdog.log	watchdog run log
update-service	omm_upd_server.log	UPDServer run log file
	omm_upd_agent.log	UPDAgent run log file
	update-manager.log	UPDManager run log file
	install.log	Installation log file of the upgrade service
	uninstall.log	Uninstallation log file of the upgrade service

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Log File Name	Description
	catalina.<*Time*>.log, catalina.out, host-manager.<*Time*>.log, localhost.<*Time*>.log, manager.<*Time*>.log, manager_access_log.<*Time*>.txt, web_service_access_log.<*Time*>.txt, catalina.log, gc-update-service.log.0.current, update-manager.controller, update-web-service.controller, update-web-service.log, commit_rm_distributed.log, commit_rm_upload_package.log, common_omagent_operator.log, forbid_monitor.log, initialize_package_atoms.log, initialize_unzip_pack.log, omm-upd.log, register_patch_pack.log, resume_monitor.logrollback_clear_patch.log, unregister_patch_pack.log, update-rcommupd.log, update-rcupdatemanager.log, and update-service.log	Run log file of the upgrade service
upgrade	upgrade.log_<*Time*>	OMS upgrade log file
	rollback.log_<*Time*>	OMS rollback log file
sudo	sudo.log	Sudo script execution log

Log levels

Table 9-82 describes the log levels provided by Manager. The priority of log levels, from highest to lowest, is FATAL, ERROR, WARN, INFO, and DEBUG. Logs whose levels are higher than or equal to the set level are printed by the program. The number of printed logs decreases as the set log level increases.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-82 Log Levels

Level	Description
FATAL	Logs of this level record fatal error information about the current event handling that may result in a system crash.
ERROR	Logs of this level record error information about the current event handling that indicates a system malfunction.
WARN	Logs of this level record abnormal information about the current event handling that does not cause the system to malfunction.
INFO	Logs of this level record normal status information about the system and events.
DEBUG	Logs of this level record system and its debugging information.

Log format

Manager log formats are as follows:

Table 9-83 Log formats

Type	Component	Format	Example
Controller, HTTPd, Logman, NodeAgent, oKerberos, oldapserver, OMM, Tomcat, and upgrade	Controller, HTTPd, Logman, NodeAgent, oKerberos, oldapserver, OMM, Tomcat, and upgrade	<*yyyy-MM-dd HH:mm:ss,SSS*>|<*Log level*>|<*Name of the thread that generates the log*>|<*Message in the log*>|<*Location where the log event occurs*>	2020-06-30 00:37:09,067 INFO [pool-1-thread-1] Completed Discovering Node. com.xxx.hadoop.om.controller.tasks.nodesetup.DiscoverNodeTask.execute(DiscoverNodeTask.java:299)

9.8.2 Viewing MRS Operation Logs

The MRS console records user operations on MRS clusters and jobs. These logs are used to locate cluster issues, enabling timely resolution.

The MRS console records the following types of operations:

- Cluster operations
  - Create, delete, scale out, and scale in a cluster on the management console.
  - Create a directory and delete a directory and its files on the management console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Data operations: IAM user tasks, adding user, and adding user group.

When you create, stop, or delete a job in an MRS cluster on the management console or using the APIs, the operation will not be recorded on the **Operation Logs** page. You can check the operation records on the CTS console.

Viewing MRS Operation Logs

Step 1 Log in to the MRS console.

Step 2 In the navigation pane on the left, choose **Operation Logs**.

Operation logs in the log list are sorted in chronological order by default, with the most recent logs displayed at the top.

Table 9-84 describes various fields in a log.

Table 9-84 Log description

Parameter	Description
Operation Type	Various types of operations, including: • Cluster operations • Job operations • Data operations
Operation IP	IP address where an operation is performed. If an MRS cluster fails to be deployed, the cluster is automatically deleted, and the operation logs of the automatically deleted cluster do not contain the **Operation IP** of the user.
Operation	Operation details. The value can contain a maximum of 2048 characters.
Time	Operation time. For a deleted cluster, only logs generated within the last six months are displayed. To view logs generated six months ago, contact technical support.
Enterprise Project	Enterprise project to which the cluster belongs

----End

9.8.3 Viewing MRS Cluster History

On the MRS management console, you can view terminated or unsubscribed MRS clusters under your account. This allows you to check the status of MRS cluster creation.

Viewing Information of a Historical Cluster

Step 1 Log in to the MRS console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 In the navigation pane on the left, choose **Clusters** > **Cluster History**. On this page, you can view terminated clusters.

Click the name of a terminated cluster. On the displayed page, you can view the cluster configuration, node, auto scaling, component, job, bootstrap action, and tag information.

Table 9-85 Basic cluster information

Parameter	Description
Cluster Name	Name of a cluster. The cluster name is set when the cluster is created.
Cluster Status	Cluster status information.
Cluster Version	Cluster version.
Cluster Type	Type of a cluster.
Cluster ID	Unique identifier of a cluster, which is automatically assigned when a cluster is created.
Created	Time when a cluster is created.
AZ	Availability zone (AZ) in the region of a cluster, which is set when a cluster is created.
Default Subnet	Subnet selected during cluster creation. A subnet provides dedicated network resources that are isolated from other networks to enhance network security.
VPC	VPC selected during cluster creation. A VPC is a secure, isolated, and logical network environment.
OBS Permission Control	Mapping between MRS users and OBS permissions.
Data Connection	Type of the data connection associated with the cluster.
Agency	Agency bound to or modified by the cluster.
Key Pair	Name of a key pair, which is set during cluster creation. This parameter is not displayed if password is used as the login mode.
Kerberos Authentication	Whether Kerberos authentication is enabled for logging in to Manager. During MRS cluster creation, carefully set the Kerberos authentication parameter. Once the cluster is created, Kerberos authentication cannot be manually enabled or disabled. To change the Kerberos authentication settings, you must create a new MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Enterprise Project	Enterprise project to which a cluster belongs. You can only click the name of an enterprise project on the **Active Clusters** page to access the **Enterprise Project Management** page.
Security Group	Security group name of a cluster.
Streaming Core Node LVM	Whether the Logical Volume Manager (LVM) function is enabled for streaming Core nodes.
Data Disk Key Name	Name of the key used to encrypt data disks. The used keys can be managed on the KMS console.
Data Disk Key ID	ID of the key used to encrypt data disks.
Component Version	Version of each component installed in the cluster.
Agency	Binding an agency to the cluster allows the ECSs or BMSs in the cluster to manage specific resources by granting them the necessary permissions.

----End

9.8.4 Viewing MRS Cluster Audit Logs

The **Audit** page records user operations on Manager. Administrators can view user operation records on Manager. For details, see **Audit logs**.

This section describes how to view and export audit logs on MRS Manager for post-event tracing, fault cause locating, and responsibility division of security events.

Viewing Audit Logs (MRS 3.*x* or Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Audit**. The **Audit** page displays audit information of FusionInsight Manager, including the operation type, risk level, start time, end time, user, host, service, instance, and operation result.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-86 Audit information

Parameter	Description
Operation Type	Type of operation performed on FusionInsight Manager. The options include **User_Manager**, **Cluster**, **Service**, **Alarm**, **Backup and Restoration**, and **Tenant**.
Severity	Security level of each audit log. The options include **Critical**, **Major**, **Minor**, and **Notice**.
Started	Time when the operation starts
Completed	Time when the operation ends
User	Name of the user who performs the operation
Host	Cluster node where the operation is performed
Service	Cluster service involved in the operation
Instance	Cluster role instance involved in the operation
Operation Result	User operation result. The options include **Successful**, **Failed**, and **Unknown**.

- You can select audit logs at the **Critical**, **Major**, **Minor**, or **Notice** level from the **All risk levels** drop-down list.
- In **Advanced Search**, you can set filter criteria to query audit logs.
  a. You can query audit logs by user management, cluster, service, and health in the **Operation Type** column.
  b. In the **Service** column, you can select a service to query corresponding audit logs.
     You can select **--** to search for audit logs using all search criteria except the service.
  c. You can query audit logs by operation result. The options are **All**, **Successful**, **Failed**, and **Unknown**.
- You can click to manually refresh the current page or click to choose the columns to display on the page.
- Click **Export All** to export all audit information at a time, in **TXT** or **CSV** format.

----End

Viewing Audit Logs (Versions Earlier Than MRS 3.*x*)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **Audit** to view default audit logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 If an audit log contains more than 256 characters, click expand button to view full audit details.

- Records are sorted in descending order by the **Occurred** column as the default setting. To change the sorting mode, click **Operation Type**, **Severity**, **Occurred**, **User**, **Host**, **Service**, **Instance**, or **Operation Result**.
- You can filter all alarms of the same severity, including both cleared and uncleared alarms, by **Severity**.

Exported audit log files contain the following columns:

Table 9-87 Audit log file information

Parameter	Description
Sno	Number of audit logs generated by MRS Manager. The number increases by 1 each time an audit log is generated.
Operation Type	Operation type. The options are **Alarm**, **Auditlog**, **Backup and Restoration**, **Cluster**, **Collect Log**, **Host**, **Service**, **Tenant**, and **User_Manager**. **User_Manager** is available only for clusters with Kerberos authentication enabled. Each option contains varying operation types. For example, **Alarm** includes **Export alarms**, **Cluster** includes **Start cluster**, and **Tenant** includes **Add tenant**.
Severity	Security level of each audit log, including **Critical**, **Major**, **Minor**, and **Informational**.
Start Time	Time when the operation starts
End Time	Time when the operation ends
User IP Address	IP address used by a user to perform the operation
User	Name of the user who performs the operation
Host	Node where the operation is performed. The information is not saved if the operation does not involve a node.
Service	Service on which the operation is performed. The information is not saved if the operation does not involve a service.
Instance	Role instance on which the operation is performed. The information is not saved if the operation does not involve a role instance.
Operation Result	User operation result. The options include **Successful**, **Failed**, and **Unknown**.
Content	Detailed information of the operation

Step 4 Click **Advanced Search**. In the search area, set search criteria and click **Search** to view audit logs of a specified type. Click **Reset** to clear the search criteria.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Start Time and **End Time** specify the start time and end time of the time range. You can search for alarms generated within the time range.

Step 5 To export all audit log files in the log list, click **Export All**.

Step 6 To export an audit log file, select it in the log list and click **Export**.

----End

Helpful Links

- Audit logs of MRS cluster components are classified by name and stored in the **/var/log/Bigdata/audit** directory on each node in the cluster. You can change the maximum number of audit log files that can be backed up by OMS. For details, see **Configuring the Number of Local Backups of MRS Cluster Audit Logs**.
- You can save the audit logs of Manager to other archive servers. For details, see **Configuring Dumping for MRS Cluster Audit Logs**.

9.8.5 Viewing Role Instance Logs of MRS Components

Once an MRS cluster is created, you can easily view the logs of each component role instance and download specific log files on the Manager page. This makes it easier to quickly locate and analyze any issues.

Notes and Constraints

The operations described in this section only apply to MRS 3.*x* or later.

Viewing Role Instance Logs

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

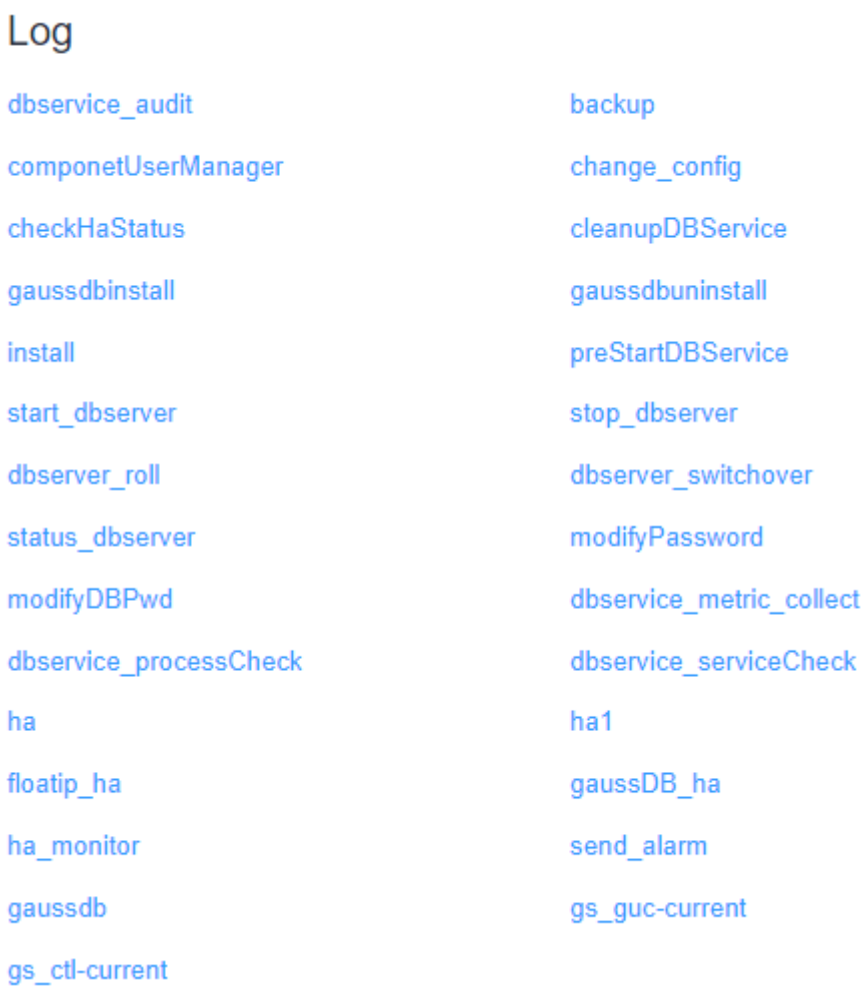
Step 2 Choose **Cluster** > **Services**, and click a service name. Then click the **Instances** tab of the service and click the name of the target instance to access the instance status page.

Step 3 In the **Log** area, click the name of a log file to preview its content online.

- On the **Hosts** page, click a host name. In the instance list of the host, you can view the log files of all role instances on the host.
- By default, a maximum of 100 lines of logs can be displayed. You can click **Load More** to view more logs. Click **Download** to download the log file to the local PC. For how to download service logs in batches, see **Downloading MRS Cluster Logs**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-1 Viewing instance logs



----End

9.8.6 Searching for MRS Cluster Logs Online

On the MRS Manager page, administrators can easily search for and view all component log files by node or component role to identify faults through keyword analysis.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Searching for Logs Online

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Log** > **Online Search**.

Step 3 Set parameters based on **Table 9-88**. You can select a log search duration or click to specify **Start Time** and **End Time**.

Table 9-88 Log search parameters

Parameter	Description
Search Content	Keywords or regular expression to be searched for
Service	Service or module for which you want to query logs
File	Log files to be searched for when only one role is selected
Lowest Log Level	Lowest level of logs to be queried. After you select a level, the logs of this level and higher levels are displayed. The levels in ascending order are as follows: TRACE < DEBUG < INFO < WARN < ERROR < FATAL
Host Scope	• Click ... to select required hosts. If no host is specified, all hosts are selected by default. • Enter the host name of the node for which you want to query logs or the IP address of the management plane. • Use commas (,) to separate IP addresses, for example, **192.168.10.10**,**192.168.10.11**. • Use hyphens (-) to indicate an IP address segment if the IP addresses are consecutive, for example, **192.168.10.[10-20]**. • Use hyphens (-) to indicate an IP address segment if the IP addresses are consecutive, and use commas (,) to separate IP address segments, for example, **192.168.10.[10-20,30-40]**. • A maximum of 10 IP address expressions can be entered at a time. A maximum of 2,000 hosts can be matched for all entered expressions at a time.
Advanced Configurations	• **Max Quantity**: maximum number of logs that can be displayed at a time. If the number of queried logs exceeds the value of this parameter, the earliest logs will be ignored. If this parameter is not set, the maximum number of logs that can be displayed at a time is not limited. • **Timeout Duration**: log query timeout duration. This parameter is used to limit the maximum log query time on each node. When the query times out, the query is stopped and the logs that have been searched for are still displayed.

Step 4 Click **Search**. **Table 9-89** describes the fields in search results.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-89 Parameters in search results

Parameter	Description
Time	Time when a line of log is generated
Host Name	Host name of the node where the log file recording the line of log is located
Location	Path of the log file recording the line of log Click the location information to go to the online log browsing page. By default, 100 lines of logs before and 100 lines after the line of log are displayed. You can click **Load More** on the top or bottom of the page to view more logs. Click **Download** to download the log file to the local PC.
Line No.	Line number of a line of log in the log file
Level	Level of the line of log
Log	Log content

You can click **Stop** to forcibly stop the search. You can view the search results in the list.

Step 5 Click **Filter** to filter the logs to display on the page. **Table 9-90** lists the fields that you can use to filter logs. After you configure these parameters, click **Filter** to search for logs meeting the search criteria. You can click **Reset** to clear the information that you have filled in.

Table 9-90 Parameters for filtering logs

Parameter	Description
Keywords	Keywords of the logs to be searched for
Host Name	Name of the host to be searched for
Location	Path of the log file to be searched for
Started	Start time for logs to be searched for
Completed	End time for logs to be searched for

----End

Helpful Links

- For details about how to export logs generated by all instances of each service role, see **Downloading MRS Cluster Logs**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For details about how to collect the stack information of a specified role or instance, see **Collecting MRS Cluster Service Stack Information**.
- For details about how to change the log level of FusionInsight Manager, see **Configuring Default Log Level and Archive File Size for MRS Components**.

9.8.7 Downloading MRS Cluster Logs

Manager allows you to export logs from all instances of each service in batches, eliminating the need to manually log in to individual nodes to obtain the logs.

Downloading MRS Cluster Logs (MRS 3.*x* or Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Log** > **Download**.

Step 3 Select a log download range.

1. **Service**: Click and select a service.
2. **Host**: Enter IP addresses of the hosts where the service is deployed. You can also click to select hosts.
3. **Max. Concurrent Nodes**: Set the maximum number of nodes whose log files can be simultaneously collected. This parameter is only available for MRS 3.3.0 or later clusters.
4. Click  in the upper right corner and set **Start Time** and **End Time**.

Step 4 Click **Download**.

The downloaded log package contains the topology information of the start time and end time, helping you quickly find the log you need.

The topology file is named in the format of **topo_<***Topology structure change time***>.txt**. The file contains the node IP address, host name, and service instances that reside on the node. (OMS nodes are identified by **Manager:Manager**.)

Example:

```
192.168.204.124|suse-124|
DBService:DBServer;KrbClient:KerberosClient;LdapClient:SlapdClient;LdapServer:SlapdServer;Manager:Manager;meta:meta
```

----End

Downloading MRS Cluster Logs (Versions Earlier Than MRS 3.*x*)

- You have obtained the access key ID (AK) and secret access key (SK) of the account.
- A parallel file system has been created in OBS.

Step 1 Log in to MRS Manager by referring to **Accessing MRS Manager** and click **System**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Click **Export Log** under **Maintenance**.

Step 3 Select a service for **Service**. Set **Host** to the IP address of the host where the service is deployed. Select the corresponding time for **Start Time** and **End Time**.

Step 4 In **Export To**, select a path for storing logs. This parameter is available only for clusters with Kerberos authentication enabled.

- **Local PC**: log files are saved to the local environment. Go to **Step 8**.
- **OBS**: logs are saved to OBS. This is the default option. Go to **Step 5**.

Step 5 Set **OBS Path** to the OBS path for storing service log files.

Enter a full path that does not start with a slash (/) and is no more than 900 bytes. The system will automatically create the path if it does not exist.

Step 6 In **Bucket**, enter the name of the created OBS file system.

Step 7 Set **AK** and **SK** to the user's access key ID and secret access key.

Step 8 Click **OK**.

----End

9.8.8 Collecting MRS Cluster Service Stack Information

On Manager, administrators can collect stack information of a specified role or instance and download it to a local directory to meet service needs.

The following information can be collected:

1. jstack information.
2. jmap -histo information.
3. jmap -dump information.
4. Thr jstack and jmap-histo information can be collected continuously for comparison.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Collecting Stack Information

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > *Name of the service whose stack information is to be collected*.

Step 3 On the displayed page, Choose **More** > **Collect Stack Information**.

- To collect stack information of multiple instances, go to the instance list, select the desired instances, and choose **More** > **Collect Stack Information**.
- To collect stack information of a single instance, click the desired instance and choose **More** > **Collect Stack Information**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 In the displayed dialog box, select the desired role and content, configure advanced options (retain the default settings if there is no special requirement), and click **OK**.

Step 5 After the collection is successful, click **Download**.

----End

Downloading Stack Information

Step 1 Click **Cluster**, click the name of the desired cluster, click **Services**, and click the target service. Choose **More** > **Download Stack Information** in the upper right corner.

Step 2 Select the desired role and content and click **Download** to download the stack information to the local PC.

----End

Clearing Stack Information

Step 1 Click **Cluster**, click the name of the desired cluster, click **Services**, and click the target service.

Step 2 Choose **More** > **Clear Stack Information** in the upper right corner.

Step 3 Select the desired role and content and configure **File Directory**. Click **OK**.

----End

9.8.9 Configuring Default Log Level and Archive File Size for MRS Components

You can change the log levels of FusionInsight Manager. For a specific service, you can change the log level and the log file size to prevent the failure in saving logs due to insufficient disk space.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Impact on the System

The services need to be restarted for the new configuration to take effect. During the restart, the services are unavailable.

Changing the Manager Log Level

1. Log in to the active management node as user **omm**.
2. Run the following command to switch to the required directory:
   cd ${BIGDATA_HOME}/om-server/om/sbin
3. Run the following command to change the log level:
   ./setLogLevel.sh*Log level parameters*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The priorities of log levels are FATAL, ERROR, WARN, INFO, and DEBUG in descending order. Logs whose levels are higher than or equal to the set level are printed. The number of printed logs decreases as the configured log level increases.

- **DEFAULT**: After this parameter is set, the default log level is used.
- **FATAL**: critical error log level. After this parameter is set, only logs of the **FATAL** level are printed.
- **ERROR**: error log level. After this parameter is set, logs of the **ERROR** and **FATAL** levels are printed.
- **WARN**: warning log level. After this parameter is set, logs of the **WARN**, **ERROR**, and **FATAL** levels are printed.
- **INFO** (default): informational log level. After this parameter is set, logs of the **INFO**, **WARN**, **ERROR**, and **FATAL** levels are printed.
- **DEBUG**: debugging log level. After this parameter is set, logs of the **DEBUG**, **INFO**, **WARN**, **ERROR**, and **FATAL** levels are printed.
- **TRACE**: tracing log level. After this parameter is set, logs of the **TRACE**, **DEBUG**, **INFO**, **WARN**, **ERROR**, and **FATAL** levels are printed.

The log levels of components are different from those defined in open-source code.

4. Download and view logs to verify that the log level settings have taken effect. For details, see **Downloading MRS Cluster Logs (MRS 3.x or Later)**.

Changing the Service Log Level and Log File Size

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services**.

Step 3 Click a service in the service list. On the displayed page, click the **Configuration** page.

Step 4 On the displayed page, click the **All Configuration** tab. Expand the role instance displayed on the left of the page. Click **Log** of the role to be modified.

Step 5 Search for each parameter and obtain the parameter description. On the parameter configuration page, select the required log level or change the log file size. The unit of the log file size is MB.

- The system automatically deletes logs based on the configured log size. To save more logs, set a larger log file size. To ensure the integrity of log files, you are advised to manually back up the log files to another directory based on the actual service volume before the log files are cleared based on clearance rules.
- The log level of some services cannot be changed on the console.
- The log level and log file size of KrbServer, LdapServer, and DBService cannot be changed.

Step 6 Click **Save**. In the **Save Configuration** dialog box, click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Download and view logs to verify that the log level settings have taken effect.

----End

9.8.10 Configuring the Number of Local Backups of MRS Cluster Audit Logs

Audit logs of cluster components are classified by name and stored in the **/var/log/Bigdata/audit** directory on each cluster node. The OMS automatically backs up the audit log directories at 03:00 every day.

The audit log directory on each node is compressed and named in the *<Node IP address>***.tar.gz** format. All compressed files are compressed and named in the *<yyyy-MM-dd_HH-mm-ss>***.tar.gz** format and saved in the **/var/log/Bigdata/audit/bk/** directory on the active management node. In addition, the standby management node saves a copy of the file.

By default, a maximum of 90 OMS backup files can be retained. This section describes how to configure the maximum number.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Configuring the Number of Local Backups of MRS Cluster Audit Logs

Step 1 Log in to the active OMS node in the cluster as user **root**.

WARNING

Change the number of audit log backup files only on the active OMS node. If you change it on the standby OMS node, the cluster may fail to work properly.

Step 2 Run the following command to switch to user **omm**:

```
su - omm
```

Step 3 Run the following command to switch to the required directory:

```
cd ${BIGDATA_HOME}/om-server/om/sbin
```

Step 4 Run the following command to change the maximum number of audit log backup files to be retained:

```
./modifyLogConfig.sh -m Maximum number of audit log backup files that can be retained
```

The default value is **90**. The value ranges from **0** to **365**. A larger value means to consume more disk space.

If the following information is displayed, the operation is successful:

```
Modify log config successfully
```

----End

9.8.11 Configuring Dumping for MRS Cluster Audit Logs

The audit logs of FusionInsight Manager are stored in the database by default. If the audit logs are retained for a long time, the disk space of the data directory

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

may be insufficient. To store audit logs to another archive server, administrators can set the required dump parameters to automatically dump these logs. This facilitates the management of audit logs.

If you do not configure the audit log dumping, the system automatically saves the audit logs to a file when the number of audit logs reaches 100,000 pieces. The path is **${BIGDATA_DATA_HOME}/dbdata_om/dumpData/iam/operatelog** on the active management node. The file name format is **OperateLog_store_***YY_MM_DD_HH_MM_SS***.csv**. A maximum of 50 historical audit log files can be saved.

Archived audit logs will not be displayed on FusionInsight Manager. Instead, only new audit logs generated after the old logs are automatically saved are displayed.

Configuring Audit Log Dumping to an SFTP Server (MRS 3.*x* or Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Audit** > **Configuration**.

Step 3 Click the switch on the right of **Audit Log Dumping Flag**.

Audit Log Dump is disabled by default. If  is displayed, **Audit Log Dump** is enabled.

Step 4 Set dumping parameters based on **Table 9-91**.

Table 9-91 Audit log dump parameters

Parameter	Example Value	Description
SFTP IP Mode	IPv4	Mode of the destination IP address. The value can be **IPv4** or **IPv6**.
SFTP IP	192.168.*xxx.xxx*	SFTP server for storing dumped audit logs. You are advised to use the SFTP service based on SSH v2 to prevent security risks.
SFTP Port	22	Connection port of the SFTP server for storing dumped audit logs
Save Path	/opt/omm/oms/auditLog	Path for storing audit logs on the SFTP server
SFTP Username	root	User name for logging in to the SFTP server
SFTP Password	-	Password for logging in to the SFTP server
SFTP Public Key	-	(Optional) Public key of the SFTP server. Set this parameter to the public key of the SFTP server; or, there may be security risks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Example Value	Description
Dumping Mode	By Quantity	Dump mode. Value options are as follows: • **By Quantity**: If the number of pieces of logs reaches the value of this parameter (**100000** by default), the logs are dumped. • **By Time**: specifies the date when logs are dumped. The dumping frequency is once a year.
Dumping Date	11-06	This parameter is available only when **Dumping Mode** is set to **By time**. After you select a dump date, the system starts dumping on this date. The logs to be dumped include all the audit logs generated before January 1 00:00 of the current year.

If the SFTP public key is empty, the system displays a security risk warning. Evaluate the security risk and then save the configuration.

Step 5 Click **OK** to complete the settings.

Key fields in the audit log dump file are as follows:

- **USERTYPE** indicates the user type. Value **0** indicates the **Human-machine** user, and value **1** indicates the **Machine-machine** user.
- **LOGLEVEL** indicates the security level. Value **0** indicates Critical, value **1** indicates Major, value **2** indicates Minor, and value **3** indicates Warning.
- **OPERATERESULT** indicates the operation result. Value **0** indicates that the operation succeeded, and value **1** indicates that the operation failed.

----End

Configuring Audit Log Dumping to OBS (MRS Versions Earlier Than 3.*x*)

Storing MRS audit logs in the system may result in insufficient disk space of the data directory. To facilitate audit log management, you can enable automatic export of audit logs to a specified directory on the OBS server.

Audit logs exported to the OBS server include service audit logs and management audit logs.

- Service audit logs are automatically compressed and stored in the **/var/log/Bigdata/audit/bk/** directory on the active management node at 03:00:00 every day. Log files are named in the format of *<yyyy-MM-dd_HH-mm-ss>***.tar.gz**. By default, only log files generated within the past seven days are stored, while those generated beyond seven days are automatically deleted.
- Management audit logs are exported to OBS starting from the date of the last export until the current execution date of the export task. When there are 100,000 entries in a management audit log file, the system automatically dumps the first 90,000 entries to a local file and keeps the remaining 10,000

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

in the database. Dumped log files are saved in the **${BIGDATA_DATA_HOME}/dbdata_om/dumpData/iam/operatelog** directory of the active management node. These files are named in the format of **OperateLog_store**_*YY_MM_DD_HH_MM_SS*.**csv**. A maximum of 50 historical audit log files can be saved.

Prerequisites

- You have obtained the access key ID (AK) and secret access key (SK) of the account.
- A parallel file system has been created in OBS.

Procedure

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 Choose **Export Audit Log** under **Maintenance**.

Table 9-92 Parameters for exporting audit logs

Parameter	Example Value	Description
Export Audit Log	On	(Mandatory) Whether to enable audit log export. The options are: • **On**: Audit log export is enabled. • **Off**: Audit log export is disabled.
Started	07/24/2017 09:00:00	(Mandatory) Start time for exporting audit logs.
Period	1 day	(Mandatory) Interval for exporting audit logs. The interval ranges from 1 to 5 days.
Bucket	mrs-bucket	(Mandatory) Name of the OBS file system to which audit logs are exported.
OBS Path	opt/omm/oms/auditLog	(Mandatory) OBS path to which audit logs are exported.
AK	-	(Mandatory) User's access key ID.
SK	-	(Mandatory) User's secret access key.

Service audit log files and management audit log files are stored in **service_auditlog** and **manager_auditlog** on OBS, respectively.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.9 MRS Cluster Security Configuration

9.9.1 Cluster Mutual Trust Management

9.9.1.1 Overview of Mutual Trust Between MRS Clusters

Function Description

By default, users of a big data cluster in security mode can only access resources in the cluster but cannot perform identity authentication or access resources in other clusters in security mode.

Concepts of Cluster Mutual Trust

- **Domain**

  The secure usage scope of users in each system is referred to as a domain. Each MRS Manager must have a unique domain name. Cross-Manager access allows users to use resources across domains.

- **User Encryption**

  Mutual trust can be configured across FusionInsight Managers. The current Kerberos server supports only the aes256-cts-hmac-sha1-96:normal and aes128-cts-hmac-sha1-96:normal encryption types for encrypting cross-domain users, and the encryption types cannot be changed.

- **User Authentication**

  After cross-Manager mutual trust is configured, if a user with the same name exists in two systems and the user in the peer system has the permission to access a resource in that system, this user can also access the remote resource.

- **Direct Mutual Trust**

  The system saves the mutual trust ticket of the peer system in two clusters with mutual trust configured and uses the mutual trust ticket to access the peer system.

9.9.1.2 Changing the System Domain Name of an MRS Cluster

Scenarios

The secure usage scope of users in each system is called a domain. Each system must have a unique domain name. The domain name of FusionInsight Manager is generated during installation. The system administrator can change the domain name on FusionInsight Manager.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

WARNING

Changing the system domain name is a high-risk operation. Before performing operations in this section, ensure that the OMS data has been backed up by referring to **Backing Up Manager Data (MRS 3.x and Later Versions)**.

Notes and Constraints

This section applies only to MRS 3.1.0 or later.

Impact on the System

- During the configuration process, all clusters must be restarted and will be unavailable while the restart is in progress.
- After the domain name is changed, the passwords of the Kerberos administrator and OMS Kerberos administrator will be initialized. You need to use the default passwords and then change the passwords. If a component user whose password is generated randomly by the system is used for identity authentication, see **Downloading MRS Cluster User Credentials** to download the keytab file again.
- After the domain name is changed, passwords of the **admin** user, component user, and human-machine user added by the system administrator before the domain name change will be reset to the same one. Change these passwords. The reset password consists of two parts: a system- generated part and a user- defined part. The system- generated part is **Admin@123**, which is the default password. For details about the user-defined part, see description of **Password Suffix** in **Table 9-94**. For example, if the system generates **Admin@123** and the user sets **Test#$%@123**, the new password after reset is **Admin@123Test#$%@123**.
- The new password must meet the password policies. To obtain the new human-machine user password, log in to the active OMS as user **omm** and run the following script:

  sh ${BIGDATA_HOME}/om-server/om/sbin/get_reset_pwd.sh *Password suffix user_name*

  - *Password suffix* is a parameter set by the user. If it is not specified, the default value **Admin@123** is used.
  - *user_name* is optional. The default value is **admin**.
  - Commands carrying authentication passwords pose security risks. Disable historical command recording before running such commands to prevent information leakage.

  Example:

  sh ${BIGDATA_HOME}/om-server/om/sbin/get_reset_pwd.sh Test#$%@123

  ```
  To get the reset password after changing cluster domain name.
    pwd_min_len     : 8
    pwd_char_types  : 4
  The password reset after changing cluster domain name is: "Admin@123Test#$%@123"
  ```

  In this example, **pwd_min_len** and **pwd_char_types** indicate the minimum password length and number of password character types respectively defined

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

in the password policies. **Admin@123Test#$%@123** indicates the human-machine user password after the system domain name is changed.

- After the system domain name is changed, the reset password consists of two parts: a system- generated part and a user- defined part. The new password must meet the password policies. If the password is not long enough, one or multiple at signs (@) are added between **Admin@123** and the user-defined part. If there are five character types, a space is added after **Admin@123**.

  When the user-defined part is **Test@123** and the default user password policy is used, the new password is **Admin@123Test@123**. The password contains 17 characters of four types. To meet the current password policy, the new password is processed according to **Table 9-93**.

  Table 9-93 Password processing

Minimum Password Length	Number of Character Types	Processing Against the Password Policy	New Password
8 to 17 characters	4	The user password policy is met.	Admin@123Test@123
18 characters	4	Add an at sign (@).	Admin@123@Test@123
19 characters	4	Add two at signs (@).	Admin@123@@Test@123
8 to 18 characters	5	Add a space.	Admin@123 Test@123
19 characters	5	Add a space and an at sign (@).	Admin@123 @Test@123
20 characters	5	Add a space and two at signs (@).	Admin@123 @@Test@123

- After the system domain name is changed, download the **keytab** file for the machine-machine user added by the system administrator before the domain name is changed.
- After the system domain name is changed, download and install the client again.

Prerequisites

- The system administrator has clarified service requirements and planned domain names for the systems.

  A domain name can contain only uppercase letters, numbers, periods (.), and underscores (_), and must start with a letter or number.
- The running status of all components in the Manager clusters is **Normal**.
- The **acl.compare.shortName** parameter of the ZooKeeper service of all clusters in Manager is set to default value **true**. Otherwise, change the value to **true** and restart the ZooKeeper service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Changing the System Domain Name

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **Domain and Mutual Trust**.

Step 3 Modify required parameters.

Table 9-94 Related parameters

Parameter	Example Value	Description
Local Domain	**MRSTEST.COM**	Planned domain name of the system.
Password Suffix	**Admin@123**	Part of the password set by the user after the password of the human-machine user is reset. This parameter is mandatory. The default value is **Admin@123**. This parameter takes effect only after **Local Domain** is modified. The following requirements must be met: • The password contains 8 to 16 characters. • The password contains at least three types of the following: uppercase letters, lowercase letters, numbers, and the following special characters: `~!@#$%^&*()-_=+|[{}];:',<.>/? and space.

Step 4 Determine whether to set up mutual trust between managers across cluster.

- If yes, perform operations by referring to **Configuring Mutual Trust Between MRS Clusters** and skip subsequent steps in this section.
- If no, go to **Step 5**.

Step 5 Click **OK**. Proceed with the subsequent steps only after the modification is complete.

Step 6 Log in to the active management node as user **omm**.

Step 7 Run the following command to update the domain configuration:

sh ${BIGDATA_HOME}/om-server/om/sbin/restart-RealmConfig.sh

The command is executed successfully if the following information is displayed:

```
Modify realm successfully. Use the new password to log in to FusionInsight again.
```

After the restart, some hosts and services may be inaccessible and an alarm is generated. This issue can be automatically resolved in about 1 minute after **restart-RealmConfig.sh** is executed.

Step 8 Log in to FusionInsight Manager using the new password of user **admin** (for example, **Admin@123Admin@123**). On the dashboard page, click or **More** and select **Restart**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In the displayed dialog box, enter the password of the current login user and click **OK**.

In the displayed dialog box, click **OK**. Wait for a while until a message indicating that the operation is successful is displayed. Click **Finish**.

Step 9 Log out of FusionInsight Manager and then log in again. If the login is successful, the configuration is successful.

Step 10 Log in to the active management node as user **omm** and run the following command to update the configurations of the job submission client:

sh /opt/executor/bin/refresh-client-config.sh

----End

9.9.1.3 Configuring Mutual Trust Between MRS Clusters

When clusters in different Manager systems with security modes need to access each other's resources, system administrators can establish mutual trust between the systems, allowing external system users to access resources in the local system.

Without configuring cross-cluster trust, resources in each cluster can only be accessed by local users. The secure usage scope of users in each system is referred to as a domain. Each MRS Manager must have a unique domain name. Cross-Manager access allows users to use resources across domains.

Notes and Constraints

In MRS 3.*x* or later, a maximum of 500 mutual-trust clusters can be configured for a cluster.

Impact on the System

- Once cross-Manager cluster mutual trust is configured, users from an external system can be utilized in the local system. The system administrator should regularly review user permissions in Manager to ensure they align with enterprise service and security requirements.
- If you set up mutual trust between clusters, affected services need to be restarted and services will be interrupted.
- After cross-Manager cluster mutual trust is configured, internal Kerberos users **krbtgt/***Local cluster domain name***@***External cluster domain name* and **krbtgt/***External cluster domain name***@***Local cluster domain name* are added to the two mutually trusted clusters. The internal users cannot be deleted.
  - For MRS 2.*x* or earlier, the default password is **Crossrealm@123**.
  - For MRS 3.*x* or later, the system administrator should regularly update the passwords of the four users in the mutually trusted systems and ensure they are consistent. For details, see **Changing the Passwords for MRS Cluster Component Running Users**. When the passwords are changed, the connectivity between cross-cluster service applications may be affected.
- For MRS 3.*x* or later, after cross-Manager cluster mutual trust is configured, the clients of each cluster need to be redownloaded and reinstalled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- After cross-Manager cluster mutual trust is configured, you need to check whether the system works properly and how to access resources of the peer system as a user of the local system. For details, see **Configuring User Permissions for Mutually Trusted MRS Clusters**.

Prerequisites

- The system administrator has clarified service requirements and planned domain names for the systems. A domain name can contain uppercase letters, numbers, periods (.), and underscores (_), and must start with a letter or number.
- The domain names of the two Managers are different. When an ECS or BMS cluster is created on MRS, a unique system domain name is randomly generated. Generally, you do not need to change the system domain name.
- The two clusters do not have the same host name or the same IP address.
- The system time of the two clusters is consistent, and the NTP services in the two systems use the same clock source.
- If the two clusters for which mutual trust is configured are MRS 3.5.0-LTS or later, ensure that the encryption algorithms and modes used by the two clusters are the same.

  You can run the **cat ${CONTROLLER_HOME}/inst/conf/oms-config.ini | grep "krb5_supported_enctypes"** command in the MRS cluster and determine the encryption type based on the value of **krb5_supported_enctypes**.
  - If this parameter is left blank, the default encryption algorithm and mode are used, that is, AES256-CTS-HMAC-SHA1-96 AES128-CTS-HMAC-SHA1-96.
  - **aes256-sha1,aes128-sha1** indicates that the encryption algorithm and mode are AES256-CTS-HMAC-SHA1-96 AES128-CTS-HMAC-SHA1-96.
  - **aes256-sha2,aes128-sha2** indicates that the encryption algorithm and mode are AES256-CTS-HMAC-SHA384-192 AES128-CTS-HMAC-SHA256-128.
- The running status of all components in the Manager clusters is **Normal**.
- For MRS 3.*x* or later, the **acl.compare.shortName** parameter for the ZooKeeper service in all clusters within Manager should be set to the default value of **true**. Otherwise, change the value to **true** and restart the ZooKeeper service.

Configuring Mutual Trust Between MRS Clusters (MRS 3.*x* or Later)

Step 1 Log in to the MRS Manager of either cluster.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **Domain and Mutual Trust**.

Step 3 Modify **Peer Mutual Trust Domain**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-95 Parameters

Parameter	Example Value	Description
realm_name	**MRSTEST.COM**	Enter the domain name of the peer system.
ip_port	**10.0.0.1:21732,10.0.0.2:21732**	Enter the KDC address of the peer system. Value format: *IP address of the node accommodating the Kerberos service in the peer system:Port number* ● In dual-plane networking, enter the service plane IP address. ● If an IPv6 address is used, the IP address must be enclosed in square brackets ([]). ● Use commas (,) to separate the KDC addresses if the active and standby Kerberos services are deployed or multiple clusters in the peer system need to establish mutual trust with the local system. ● You can obtain the port number from the **kdc_ports** parameter of the KrbServer service. The default value is **21732**. To obtain the IP address of the node where the service is deployed, click the **Instance** tab on the KrbServer page and view **Service IP Address** of the KerberosServer role. For example, if the Kerberos service is deployed on nodes at **10.0.0.1** and **10.0.0.2** that have established mutual trust with the local system, the parameter value is **10.0.0.1:21732,10.0.0.2:21732**.

If you need to configure mutual trust for multiple MRS Manager systems, click + to add a new item and set parameters. Click — to delete unnecessary configurations.

Step 4 Click **OK**.

Step 5 Log in to the active management node as user **omm** and run the following command to update the domain configuration:

sh ${BIGDATA_HOME}/om-server/om/sbin/restart-RealmConfig.sh

The command is executed successfully if the following information is displayed:

```
Modify realm successfully. Use the new password to log in to FusionInsight again.
```

After the restart, some hosts and services cannot be accessed and an alarm is generated. This problem can be automatically resolved in about 1 minute after **restart-RealmConfig.sh** is run.

Step 6 Log in to FusionInsight Manager and restart the cluster or configure expired instances:

Check whether the system domain name of Manager is changed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the system domain name is changed, click ··· or **More** on the home page, click **Start**, enter the password, select the checkbox for confirming the impact, and click **OK**. Wait until the cluster is started successfully.
- If the system domain name is not changed, click ··· or **More** on the home page and select **Restart Configuration-Expired Instances**. Enter the password, select the checkbox for confirming the impact, and click **OK**. Wait until the service is restarted.

  If the Doris service is installed in the cluster, choose **Cluster** > **Services** > **Doris**. On the **Dashboard** page, choose **More** > **Restart Service** in the upper right corner, enter the password, and click **OK**. Wait until the Doris service restarts.

WARNING

Restarting a cluster or a role instance in a cluster will interrupt services. Perform this operation during off-peak hours or after confirming that the impact on upper-layer services is limited.

Step 7 Log out of FusionInsight Manager and then log in again. If the login is successful, the configuration is successful.

Step 8 Log in to FusionInsight Manager of another cluster and repeat the preceding steps.

----End

Configuring Mutual Trust Between MRS Clusters (MRS 2.*x* or Earlier)

Step 1 On the MRS management console, query all security groups of the two clusters.

- If the security groups of the two clusters are the same, go to **Step 3**.
- If the security groups of the two clusters are different, go to **Step 2**.

Step 2 On the VPC management console, choose **Access Control** > **Security Groups**. On the **Security Groups** page, locate the row containing the target security group, click **Manage Rule** in the **Operation** column.

On the **Inbound Rules** tab page, click **Add Rule**. In the **Add Inbound Rule** dialog box that is displayed, configure related parameters.

- **Priority**: The value ranges from 1 to 100. The default value is **1**, which indicates the highest priority. A smaller value indicates a higher priority.
- **Action**: Select **Allow**.
- **Protocol & Port**: Choose **Protocols** > **All**.
- **Type**: Select **IPv4** or **IPv6**.
- **Source**: Select **Security group** and the security group of the peer cluster.
  - To add an inbound rule to the security group of cluster A, set **Source** to **Security group** and the security group of cluster B (peer cluster).
  - To add an inbound rule to the security group of cluster B, set **Source** to **Security group** and the security group of cluster A (peer cluster).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For a normal cluster with Kerberos authentication disabled, perform step **Step 1** to **Step 2** to configure cross-cluster mutual trust. For a security cluster with Kerberos authentication enabled, after completing the preceding steps, continue with the subsequent steps for configuration.

Step 3 Log in to MRS Manager of the two clusters separately. Click **Service** and check whether the **Health Status** of all components is **Good**.

- If yes, go to **Step 4**.
- If no, contact O&M personnel for troubleshooting.

Step 4 Query configuration information.

1. On MRS Manager of the two clusters, choose **Services** > **KrbServer** > **Instance**. Query the **OM IP Address** of the two KerberosServer hosts.
2. Click **Service Configuration**. Set **Type** to **All**. Choose **KerberosServer** > **Port** in the navigation tree on the left. Query the value of **kdc_ports**. The default value is **21732**.
3. Click **Realm** and query the value of **default_realm**.

Step 5 On MRS Manager of either cluster, modify the **peer_realms** parameter.

Table 9-96 Parameter description

Parameter	Example Value	Description
realm_name	**MRSTEST.COM**	Domain name of the mutual-trust cluster, that is, the value of **default_realm** obtained in step **4**.
ip_port	**10.0.0.1:21732,10.0.0.2:21732**	KDC address of the peer cluster. Format: *IP address of a KerberosServer node in the peer cluster:kdc_port* The addresses of the two KerberosServer nodes are separated by a comma. For example, if the IP addresses of the KerberosServer nodes are 10.0.0.1 and 10.0.0.2 respectively, the value of this parameter is **10.0.0.1:21732,10.0.0.2:21732**.

- To configure mutual trust for multiple clusters, click  to add a new item and set parameters. To delete an item, click .
- A cluster can establish mutual trust relationships with up to 16 clusters. By default, no trust relationship exists between clusters that are trusted by the same local cluster.

Step 6 Click **Save Configuration**. In the dialog box that is displayed, select **Restart the affected services or instances** and click **OK**. If you do not select **Restart the affected services or instances**, manually restart the affected services or instances.

After **Operation successful** is displayed, click **Finish**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

> **WARNING**
>
> **Restarting a cluster or a role instance in a cluster will interrupt services. Perform this operation during off-peak hours or after confirming that the impact on upper-layer services is limited.**

Step 7 Exit MRS Manager and log in to it again. If the login is successful, the configurations are valid.

Step 8 Log in to MRS Manager of the other cluster and repeat step **Step 5** to **Step 7**.

Step 9 Perform subsequent operations by referring to **Updating the Client Configuration of Mutually Trusted Clusters (MRS 2.x or Earlier)**.

----End

Updating the Client Configuration of Mutually Trusted Clusters (MRS 2.*x* or Earlier)

After cross-cluster mutual trust is configured, the service configuration parameters are modified on MRS Manager and the service is restarted. Therefore, you need to prepare the client configuration file again and update the client.

Scenario 1:

Cluster A and Cluster B, which are peer clusters with mutual trust established, are of the same type (for example, analysis or streaming). Update the client configuration files by referring to **Updating Client Configurations (Version 2.x or Earlier)**.

- Update the client configuration file of cluster A.
- Update the client configuration file of cluster B.

Scenario 2:

Cluster A and Cluster B, which are peer clusters with mutual trust established, are of different types. Perform the following steps to update the configuration files:

- Update the client configuration file of cluster A to cluster B.
- Update the client configuration file of cluster B to cluster A.
- Update the client configuration file of cluster A.
- Update the client configuration file of cluster B.

Step 1 Log in to MRS Manager of cluster A.

Step 2 Click **Services**, and then **Download Client**.

Step 3 Set **Client Type** to **Only configuration files**.

Step 4 Set **Download to** to **Remote host**.

Step 5 Set **Host IP Address** to the IP address of the active Master node of cluster B, **Host Port** to 22, and **Save Path** to **/tmp**.

- If the default port **22** for logging in to cluster B using SSH is changed, set **Host Port** to a new port.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The value of **Save Path** contains a maximum of 256 characters.

Step 6 Set **Login User** to **root**.

If another user is used, ensure that the user has permissions to read, write, and execute the save path.

Step 7 Select **Password** or **SSH Private Key** for **Login Mode**.

- **Password**: Enter the password of user **root** set during cluster creation.
- **SSH Private Key**: Select and upload the key file used for creating the cluster.

Step 8 Click **OK** to generate a client file.

- If the following information is displayed, the client package is saved. Click **Close**.
  Client files downloaded to the remote host successfully.
- If the following information is displayed, check the username, password, and security group configurations of the remote host. Ensure that the username and password are correct and an inbound rule of the SSH (22) port has been added to the security group of the remote host. And then, go to **Step 2** to download the client again.
  Failed to connect to the server. Please check the network connection or parameter settings.

Step 9 Log in to the ECS of cluster B using VNC. For details, see **Instances** > **Logging In to a Windows ECS > Logging In to a Windows ECS Using VNC** in the *Elastic Cloud Server User Guide*.

Step 10 Run the following command to switch to the client directory, for example, **/opt/Bigdata/client**:

cd /opt/Bigdata/client

Step 11 Run the following command to update the client configuration of cluster A to cluster B:

sh refreshConfig.sh *Client installation directory Full path of the client configuration file package*

For example, run the following command:

sh refreshConfig.sh /opt/Bigdata/client /tmp/MRS_Services_Client.tar

If the following information is displayed, the configurations have been updated successfully.

```
ReFresh components client config is complete.
Succeed to refresh components client config.
```

NOTE

You can also refer to method 2 in **Updating Client Configurations (Version 2.x or Earlier)** to perform operations in **Step 1** to **Step 11**.

Step 12 Repeat step **Step 1** to **Step 11** to update the client configuration file of cluster B to cluster A.

Step 13 Refer to **Updating Client Configurations (Version 2.x or Earlier)** to update the client configuration file of the local cluster.

- Update the client configuration file of cluster A.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Update the client configuration file of cluster B.

----End

9.9.1.4 Configuring User Permissions for Mutually Trusted MRS Clusters

Once cross-Manager cluster mutual trust is configured, grant user access permissions on FusionInsight Managers to enable them to perform service operations in the mutually trusted Managers.

Prerequisites

The mutual trust relationship has been configured between two clusters (clusters A and B). The clients of the clusters have been updated.

Configuring User Permissions for Mutually Trusted Clusters (MRS 3.*x* or Later)

Step 1 Log in to MRS Manager of the local cluster.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User** to check whether the target user exists.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Click ⌄ on the left of the target user, and check whether the permissions assigned to the user group of the user and the roles meet service requirements. If not, create a role and bind the role to the user, or modify the user group or role permissions of the user.

Step 4 Create a user required by the service operations and associate the required user group or role. For details, see **Creating a User (MRS 3.x and Later)**.

Step 5 Log in to the other FusionInsight Manager and repeat **Step 2** to **Step 4** to create a user with the same name and set permissions.

----End

Configuring User Permissions for Mutually Trusted Clusters (MRS 2.*x* or Earlier)

Step 1 Log in to MRS Manager of cluster A and choose **System** > **Manage User**. Check whether cluster A has accounts that are the same as those of cluster B.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Click ⌄ on the left side of the username to view detailed user information. Verify if the user's group and role meet the service requirements.

For example, user **admin** of cluster A has the permission to access and create files in the **/tmp** directory of cluster A. Then go to **Step 4**.

Step 3 Create the accounts in cluster A and bind the accounts to the user group and roles required by the services. Then go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Choose **Service** > **HDFS** > **Instance**. Query the **OM IP Address** of **NameNode (Active)**.

Step 5 Log in to the client of cluster B.

For example, if you updated the client on the Master2 node, you must log in to the Master2 node to use the client.

Step 6 Run the following command to access the **/tmp** directory of cluster A.

hdfs dfs -ls hdfs://192.168.6.159:9820/tmp

In the preceding command, **192.168.6.159** is the IP address of the active NameNode of cluster A; **9820** is the default port for communication between the client and the NameNode.

Step 7 Run the following command to create a file in the **/tmp** directory of cluster A:

hdfs dfs -touchz hdfs://192.168.6.159:9820/tmp/mrstest.txt

If you can find the **mrstest.txt** file in the **/tmp** directory of cluster A, the cross-cluster mutual trust is successfully configured.

----End

9.9.2 MRS Cluster Security Hardening

9.9.2.1 MRS Cluster Security Hardening Policies

Hardening Tomcat

Tomcat is hardened as follows based on open-source software during FusionInsight Manager software installation and use:

- The Tomcat version is upgraded to the official version.
- Permissions on the directories under applications are set to **500**, and the write permission on some directories is supported.
- The Tomcat installation package is automatically deleted after the system software is installed.
- The automatic deployment function is disabled for projects in application directories. Only the **web**, **cas**, and **client** projects are deployed.
- Some unused **http** methods are disabled, preventing attacks by using the **http** methods.
- The default shutdown port and command of the Tomcat server are changed to prevent hackers from shutting down the server and attacking servers and applications.
- To ensure security, the value of **maxHttpHeaderSize** is changed, which enables server administrators to control abnormal requests of clients.
- The Tomcat version description file is modified after Tomcat is installed.
- To prevent disclosure of Tomcat information, the Server attributes of Connector are modified so that attackers cannot obtain information about the server.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Permissions on files and directories of Tomcat, such as the configuration files, executable files, log directories, and temporary folders, are under control.
- Session facade recycling is disabled to prevent request leakage.
- LegacyCookieProcessor is used as CookieProcessor to prevent the leakage of sensitive data in cookies.

Hardening LDAP

LDAP is hardened as follows after a cluster is installed:

- In the LDAP configuration file, the password of the administrator account is encrypted using SHA. After the OpenLDAP is upgraded to 2.4.39 or later, data is automatically synchronized between the active and standby LDAP nodes using the SASL External mechanism, which prevents disclosure of the password.
- The LDAP service in the cluster supports the SSLv3 protocol by default, which can be used safely. When the OpenLDAP is upgraded to 2.4.39 or later, the LDAP automatically uses TLS1.0 or later to prevent unknown security risks.

Hardening JDK

- If the client process uses the AES256 encryption algorithm, JDK security hardening is required. The operations are as follows:

  Obtain the Java Cryptography Extension (JCE) package whose version matches that of JDK. The JCE package contains **local_policy.jar** and **US_export_policy.jar**. Copy the JAR files to the following directory and replace the files in the directory.

  - Linux: *JDK installation directory*/**jre/lib/security**
  - Windows: *JDK installation directory***jre\lib\security**

  NOTE

  Access the Open JDK open-source community to obtain the JCE file.

- If the client process uses the SM4 encryption algorithm, the JAR package needs to be updated.

  Obtain **SMS4JA.jar** in the *client installation directory*/**JDK/jdk/jre/lib/ext/** directory, and copy the JAR package to the following directory:

  - Linux: *JDK installation directory*/**jre/lib/ext/**
  - Windows: *JDK installation directory***jre\lib\ext**

9.9.2.2 Configuring Hadoop Data Encryption During Transmission

Configuring Security Channel Encryption for Hadoop Components

Encrypted channel is an encryption protocol of remote procedure call (RPC) in HDFS. When a user invokes an RPC, the user's login name will be transmitted to RPC through RPC head. Then RPC uses Simple Authentication and Security Layer (SASL) to determine an authorization protocol (Kerberos and DIGEST-MD5) to complete RPC authorization.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

By default, the communication between components in the big data cluster is not encrypted. However, the communication can be encrypted to prevent data interception on the Internet or an untrusted network.

To do this, you just need to set related parameters to control whether to encrypt RPC channels of all Hadoop modules. The settings take effect globally.

RPC channels in Hadoop include:

- RPC channels for clients to access HDFS
- RPC channels between HDFS modules, for example, between DataNode and NameNode
- RPC channels for clients to access YARN
- RPC channels between NodeManager and ResourceManager
- RPC channels for Spark to access YARN and HDFS
- RPC channels for MapReduce to access YARN and HDFS
- RPC channels for HBase to access HDFS

On FusionInsight Manager, choose **Cluster** > **Services** > *Target service name*. Click **Configurations** and then **All Configurations**. Enter the parameter name in the search box and **restart the corresponding service for the configuration to take effect**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-97 Parameter description

Service	Parameter	Description	Default Value
HDFS	hadoop.rpc.protection	Whether to encrypt the RPC channels of each module in Hadoop. • **privacy**: indicates that the RPC channels are encrypted, and the authentication, integrity, and privacy functions are enabled. This mode may degrade the performance. • **integrity**: indicates that the RPC channels are not encrypted, and only the authentication and integrity functions are enabled. To ensure data security, exercise caution when using this mode. • **authentication**: indicates that the RPC channels are not encrypted, and only the authentication packet is required. The integrity and privacy are not required. This mode ensures performance but has security risks. This parameter applies to the RPC channels of all modules in Hadoop. • The setting takes effect only after the service is restarted. Rolling restart is not supported. • After the setting, you need to download the client configuration file again. Otherwise, HDFS cannot provide the read and write services.	• Security mode: **privacy** • Normal mode: **authentication**
HDFS	dfs.encrypt.data.transfer	Whether to encrypt the Data Transfer (DT) channels through which the client accesses DataNodes and the data transmission channels between DataNodes. This parameter is used only when **hadoop.rpc.protection** is set to **privacy**. Note that enabling encryption by default may severely affect performance when a large amount of service data is transmitted. If data transmission encryption is configured for one cluster in a pair of trusted clusters, data transmission encryption must also be configured for the peer cluster. • **true**: indicates that the channels are encrypted. • **false**: indicates that the channels are not encrypted.	false

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Parameter	Description	Default Value
HDFS	dfs.encrypt.data.transfer.algorithm	Indicates the encryption algorithm of the HDFS data transfer channels and the channels for clients to access HDFS. This parameter is available only when **dfs.encrypt.data.transfer** is set to **true**. • **3des**: default value, which indicates that the 3DES algorithm is used to encrypt data. • **rc4**: indicates that the RC4 algorithm is used to encrypt data. This value is not recommended because it may cause security risks.	3des
HDFS	dfs.encrypt.data.transfer.cipher.suites	The password suite for data encryption. This parameter can be left empty or set to **AES/CTR/NoPadding**. If this parameter is left empty, the encryption algorithm specified by **dfs.encrypt.data.transfer.algorithm** is used to encrypt data.	AES/CTR/NoPadding
HBase	hbase.rpc.protection	Whether to encrypt HBase channels, including the RPC channels for HBase clients to access the HBase server and the RPC channels between the HMaster and RegionServer. • **privacy**: indicates that the RPC channels are encrypted, and the authentication, integrity, and privacy functions are enabled. • **integrity**: indicates that the RPC channels are not encrypted, and only the authentication and integrity functions are enabled. • **authentication**: indicates that the RPC channels are not encrypted, and only the authentication packet is required. The integrity and privacy are not required. The privacy mode encrypts transmitted content, including sensitive information such as user tokens. This mode ensures the security of transmitted information, but has a great impact on performance. Compared with the other two modes, this mode reduces the read/write performance by about 60%. Modify the configuration based on the enterprise security requirements. The configuration items on the client and server must be the same.	• Security mode: **privacy** • Normal mode: **authentication**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Setting the Maximum Number of Concurrent Web Connections

To ensure web server reliability, new connections are rejected when the number of user connections reaches a specific threshold. This prevents DDOS attacks and service unavailability caused by too many users accessing the web server at the same time.

On FusionInsight Manager, choose **Cluster** > **Services** > *Target service name*. Click **Configurations** and then **All Configurations**. Enter the parameter name in the search box and **restart the corresponding service for the configuration to take effect**.

Table 9-98 Parameter description

Service	Parameter	Description	Default Value
HDFS/Yarn	hadoop.http.server.MaxRequests	Specifies the maximum number of concurrent web connections of each component.	2000
Spark/Spark2x	spark.connection.maxRequest	Specifies the maximum number of request connections of JobHistory.	5000

9.9.2.3 Configuring Kafka Data Encryption During Transmission

Scenario

Data between the Kafka client and the broker is transmitted in plain text. The Kafka client may be deployed in an untrusted network, exposing the transmitting data to leakage and tampering risks.

This topic is available for MRS 3.x or later.

Configuring Kafka Data Encryption During Transmission

The channel between components is not encrypted by default. You can set the following parameters to enable security channel encryption.

To modify parameters, log in to FusionInsight Manager, choose **Cluster** > **Services** > **Kafka**, click **Configurations** then **All Configurations**, and enter a parameter name in the search box.

NOTE

After the configuration, restart the corresponding service for the settings to take effect.

Table 9-99 describes the parameters related to transmission encryption on the Kafka server.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-99 Parameters relevant to Kafka data encryption during transmission

Parameter	Description	Default Value
ssl.mode.enable	Indicates whether to enable the Secure Sockets Layer (SSL) protocol. If this parameter is set to **true**, services relevant to the SSL protocol are started during the broker startup.	false
security.inter.broker.protocol	Indicates communication protocol between brokers. The communication protocol can be PLAINTEXT, SSL, SASL_PLAINTEXT, or SASL_SSL.	SASL_PLAINTEXT

The SSL protocol can be configured for the server or client to encrypt transmission and communication only after **ssl.mode.enable** is set to **true** and broker enables the **SSL** and **SASL_SSL** protocols.

9.9.2.4 Encrypting Data Transmission Between the Controller and Agent

Scenario

After a cluster is installed, Controller and Agent need to communicate with each other. The Kerberos authentication is used during the communication. By default, the communication is not encrypted during the communication for the sake of cluster performance. Users who have demanding security requirements can use the method described in this section for encryption.

Notes and Constraints

This topic is available for MRS 3.x or later.

Impact on the System

- Controller and all Agents automatically restart, which interrupts FusionInsight Manager.
- The performance of management nodes deteriorates in large clusters. You are advised to enable the encryption function for clusters with a maximum of 200 nodes.

Prerequisites

You have obtained the IP addresses of the active and standby management nodes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Encrypting Data Transmission Between the Controller and Agent

Step 1 Log in to the active management node as user **omm**.

Step 2 Run the following command to disable logout upon timeout:

TMOUT=0

NOTE

After the operations in this section are complete, run the **TMOUT=***Timeout interval* command to restore the timeout interval in a timely manner. For example, **TMOUT=600** indicates that a user is logged out if the user does not perform any operation within 600 seconds.

Step 3 Run the following command to go to the related directory:

cd ${CONTROLLER_HOME}/sbin

Step 4 Run the following command to enable communication encryption:

./enableRPCEncrypt.sh -t

Run the **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** command to check whether **ResHAStatus** of the active management node Controller is **Normal** and whether you can log in to FusionInsight Manager again. If yes, the enablement is successful.

Step 5 Run the following command to disable communication encryption when necessary:

./enableRPCEncrypt.sh -f

Run the **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** command to check whether **ResHAStatus** of the active management node Controller is **Normal** and whether you can log in to FusionInsight Manager again. If yes, the enablement is successful.

----End

9.9.2.5 Configuring a Trusted IP Address to Access LDAP

Scenario

By default, the LDAP service deployed in the OMS and cluster can be accessed by any IP address. To enable the LDAP service to be accessed by only trusted IP addresses, you can configure the INPUT policy in the iptables filtering list.

Notes and Constraints

This topic is available for MRS 3.x or later.

Impact on the System

After the configuration, the LDAP service cannot be accessed by IP addresses that are not configured. Before the expansion, the added IP addresses need to be configured as trusted IP addresses.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

- You have collected the management plane IP addresses and service plane IP addresses of all nodes in the cluster and all floating IP addresses.
- You have obtained the **root** user account for all nodes in the cluster.

Configuring a Trusted IP Address to Access LDAP

Configuring trusted IP addresses for the LDAP service on the OMS

Step 1 Log in to FusionInsight Manager.

Step 2 Choose **System** > **OMS** and choose **oldap** > **Modify Configuration** to view the OMS LDAP port number, that is, the value of **LDAP Listening Port**, which is 21750 by default.

Step 3 Log in to the active management node as user **root** using the IP address of the active management node.

Step 4 Run the following command to check the INPUT policy in the iptables filtering list:

iptables -L

For example, if no rule is configured, the INPUT policy is displayed as follows:

```
Chain INPUT (policy ACCEPT)
target     prot opt source               destination
```

Step 5 Run the following command to trust all IP addresses used by the cluster. Each IP address needs to be added once.

iptables -A INPUT -s *Trusted IP address* **-p tcp --dport** *Port number* **-j ACCEPT**

For example, to configure **10.0.0.1** as a trusted IP address and enable it to access port **21750**, you need to run the following command:

iptables -A INPUT -s 10.0.0.1 -p tcp --dport 21750 -j ACCEPT

Step 6 Run the following command to untrust all IP addresses. Trusted IP addresses are not affected by this rule.

iptables -A INPUT -p tcp --dport *Port number* **-j DROP**

For example, to disable all IP addresses to access port **21750**, run the following command:

iptables -A INPUT -p tcp --dport 21750 -j DROP

Step 7 Run the following command to view the modified INPUT policy in the iptables filtering list:

iptables -L

For example, after a trusted IP address is configured, the INPUT policy is displayed as follows:

```
Chain INPUT (policy ACCEPT)
target     prot opt source               destination
ACCEPT     tcp  --  10.0.0.1             anywhere            tcp dpt:21750
DROP       tcp  --  anywhere             anywhere            tcp dpt:21750
```

Step 8 Run the following command to view the rules and rule numbers in the iptables filtering list:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

iptables -L -n --line-number

```
Chain INPUT (policy ACCEPT)
num target     prot opt source               destination
1   DROP       tcp  --  0.0.0.0/0            0.0.0.0/0           tcp dpt:21750
```

Step 9 Run the following command to delete the desired rule from the iptables filtering list based on site requirement:

iptables -D INPUT *Number of the rule to be deleted*

For example, to delete rule 1, run the following command:

iptables -D INPUT 1

Step 10 Log in to the standby management node as user **root** using the standby IP address. Repeat **Step 4** to **Step 9**.

Configuring trusted IP addresses for the LDAP service in the cluster

Step 11 Log in to FusionInsight Manager.

Step 12 Choose **Cluster** > **Services** > **LdapServer**. Click **Instance** and view the LDAP nodes.

Step 13 Go to the **Configurations** page, and view the LDAP port number of the cluster, that is, the value of **LDAP_SERVER_PORT**, which is 21780 by default.

Step 14 Log in to the LDAP node as user **root** using the LDAP service IP address.

Step 15 Run the following command to view the INPUT policy in the iptables filtering list:

iptables -L

For example, if no rule is configured, the INPUT policy is displayed as follows:

```
Chain INPUT (policy ACCEPT)
target     prot opt source               destination
```

Step 16 Run the following command to configure all IP addresses used by the cluster as trusted IP addresses. Each IP address needs to be added independently.

iptables -A INPUT -s *Trusted IP address* **-p tcp --dport** *Port number* **-j ACCEPT**

For example, to configure **10.0.0.1** as a trusted IP address and enable it to access port **21780**, you need to run the following command:

iptables -A INPUT -s 10.0.0.1 -p tcp --dport 21780 -j ACCEPT

Step 17 Run the following command to untrust all IP addresses. Trusted IP addresses are not affected by this rule.

iptables -A INPUT -p tcp --dport *Port number* **-j DROP**

For example, to disable all IP addresses to access port **21780**, run the following command:

iptables -A INPUT -p tcp --dport 21780 -j DROP

Step 18 Run the following command to view the modified INPUT policy in the iptables filtering list:

iptables -L

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, after a trusted IP address is configured, the INPUT policy is displayed as follows:

```
Chain INPUT (policy ACCEPT)
target     prot opt source               destination
ACCEPT     tcp  --  10.0.0.1            anywhere            tcp dpt:21780
DROP       tcp  --  anywhere            anywhere            tcp dpt:21780
```

Step 19 Run the following command to view the rules and rule numbers in the iptables filtering list:

iptables -L -n --line-number

```
Chain INPUT (policy ACCEPT)
num target     prot opt source               destination
1   DROP       tcp  --  0.0.0.0/0            0.0.0.0/0           tcp dpt:21780
```

Step 20 Run the following command to delete the desired rule from the iptables filtering list based on site requirement:

iptables -D INPUT *Number of the rule to be deleted*

For example, to delete rule 1, run the following command:

iptables -D INPUT 1

Step 21 Log in to the LDAP node as user **root** using the IP address of another LDAP service, and repeat **Step 15** to **Step 20**.

----End

9.9.2.6 Configuring the IP Address Whitelist for Modifying Data in an HBase Read-Only Cluster

If the Replication function is enabled for HBase in an MRS 3.x and later cluster, a protection mechanism for data modification is added on the standby HBase cluster to ensure data consistency between the active and standby clusters. Upon receiving an RPC request for data modification, the standby HBase cluster checks the permission of the user who sends the request (only HBase manage users have the modification permission). Then it checks the validity of the source IP address of the request. Only modification requests from IP addresses in the white list are accepted. The IP address white list is configured by the **hbase.replication.allowedIPs** item.

Log in to FusionInsight Manager and choose **Cluster** > **Services** > **HBase**. Click **Configurations** and search for the parameter name in **Table 9-100**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-100 Parameter description

Parameter	Description	Default Value
hbase.replication.allowedIPs	Only replication requests from specified IP addresses are allowed. Only regular expressions separated by commas (,) are supported. Each pattern can be any of the following: • Regex pattern Example: 10.18.40.*, 10.18.*, 10.18.40.11 • Range pattern (Range can be specified only in the last octet) Example: 10.18.40.[10-20] If this item is empty (default value), the white list contains only the IP address of the RegionServer of the cluster, indicating that only modification requests from the RegionServer of the standby HBase cluster are accepted.	N/A

9.9.2.7 Configuring LDAP Output Audit Logs

Users can set the audit log output level of the LDAP service and output audit logs in a specified directory, for example, **/var/log/messages**. The logs output can be used to check user activities and operation commands.

WARNING

Enabling LDAP audit log output can generate a large number of logs, impacting cluster performance. Use this feature with caution.

Notes and Constraints

This topic is available for MRS 3.x or later.

Configuring the LDAP Firewall Policy

In the cluster adopting the dual-plane networking, the LDAP is deployed on the service plane. To ensure the LDAP data security, you are advised to configure the firewall policy in the cluster to disable relevant LDAP ports.

Step 1 Log in to FusionInsight Manager.

Step 2 Choose **Cluster** > **Services** > **LdapServer** and click **Configurations**.

Step 3 Check the value of **LDAP_SERVER_PORT**, which is the service port of LdapServer.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 To ensure data security, configure the firewall policy for the whole cluster to disable the LdapServer port based on the customer's firewall environment.

----End

Enabling the LDAP Audit Log Output

Step 1 Log in to any LdapServer node.

Step 2 Run the following command to edit the **slapd.conf.consumer** file, and set the value of **loglevel** to **256** (you can run the **man slapd.conf** command on the OS to view the log level definition).

cd ${BIGDATA_HOME}/FusionInsight_BASE_8.1.0.1/install/FusionInsight-ldapserver-2.7.0/ldapserver/local/template

vi slapd.conf.consumer

```
...
pidfile         [PID_FILE_SLAPD_PID]
argsfile        [PID_FILE_SLAPD_ARGS]
loglevel   256
...
```

Step 3 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **LdapServer**. Click **More** and select **Restart Service**. In the displayed dialog box, verify the current user identity, and restart the service.

----End

9.9.2.8 Updating Encryption Keys of an MRS Cluster

When you install a cluster, the system generates an encryption key automatically to protect the security information stored in the cluster, such as database user passwords and key file access passwords. If you need to change the encryption key for any reason, such as if the original key is compromised or you want to rotate the key periodically, you can update it manually.

Impact on the System

- After a cluster key is updated, a new key is generated randomly in the cluster. This key is used to encrypt and decrypt the newly stored data. The old key is not deleted, and it is used to decrypt data encrypted using the old key. After security information is modified, for example, a database user password is changed, the new password is encrypted using the new key.
- When the key is updated, the cluster is stopped and cannot be accessed.

Prerequisites

- You have obtained the IP addresses of the active and standby management nodes.
- You have stopped the upper-layer service applications that depend on the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Updating an MRS Cluster Key (MRS 3.*x* or Later)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Overview** > **More** > **Stop** (for MRS 3.3.0 or later, choose **More** > **Stop** in the upper right corner on the **Homepage**). In the displayed dialog box, enter the password for the current login user.

Click **OK**. Wait until the system displays a message indicating that the cluster is successfully stopped.

Step 3 Log in to the active management node as user **omm**.

Step 4 Run the following command to disable logout upon timeout:

TMOUT=0

After completing these operations, run the **TMOUT=***Timeout interval* command to promptly restore the timeout setting.

For example, **TMOUT=600** indicates that a user is logged out if the user does not perform any operation within 600 seconds.

Step 5 Run the following command to switch the directory:

cd ${BIGDATA_HOME}/om-server/om/tools

Step 6 Run the following command to update the cluster key:

sh updateRootKey.sh

Enter **y** as prompted.

```
The root key update is a critical operation.
Do you want to continue?(y/n):
```

The key is updated successfully if the following information is displayed:

```
Step 4-1: The key save path is obtained successfully.
...
Step 4-4: The root key is sent successfully.
```

Step 7 On FusionInsight Manager, click **Cluster**, click the name of the desired cluster, and click **Start**. (For MRS 3.3.0 or later, choose **More** > **Start** in the upper right corner of the homepage.)

In the displayed dialog box, click **OK**. Wait until a message is displayed indicating that the startup is successful.

----End

Updating an MRS Cluster Key (MRS 2.*x* or Earlier)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Services** > **More** > **Stop Cluster**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In the displayed dialog box, select **I have read the information and understand the impact.** Click **OK**. Wait until the system displays a message indicating that the operation is successful. Click **Finish**. The cluster is stopped successfully.

Step 3 Log in to the active management node.

Step 4 Run the following commands to switch the user:

sudo su - omm

Step 5 Run the following command to disable logout upon timeout:

TMOUT=0

Step 6 Run the following command to switch the directory:

cd ${BIGDATA_HOME}/om-0.0.1/tools

Step 7 Run the following command to update the cluster key:

sh updateRootKey.sh

Enter **y** as prompted.

```
The root key update is a critical operation.
Do you want to continue?(y/n):
```

The key is updated successfully if the following information is displayed:

```
...
Step 4-1: The key save path is obtained successfully.
...
Step 4-4: The root key is sent successfully.
```

Step 8 On MRS Manager, choose **Services** > **More** > **Start Cluster**.

In the displayed dialog box, click **OK**. After **Operation successful** is displayed, click **Finish**. The cluster is started.

----End

9.9.2.9 Updating the SSH Key of User omm on MRS Cluster Nodes

Scenario

During cluster installation, the system automatically generates the SSH public key and private key for user **omm** to establish the trust relationship between nodes. After the cluster is installed, if the original keys are accidentally disclosed or new keys are used, the system administrator can perform the following operations to manually change the keys.

Notes and Constraints

This topic is available for MRS 3.x or later.

Prerequisites

- The cluster has been stopped.
- No other management operations are being performed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Updating SSH Keys for User omm

Step 1 Log in as user **omm** to the node whose SSH keys need to be replaced.

If the node is a Manager management node, run the following command on the active management node.

Step 2 Run the following command to disable logout upon timeout:

TMOUT=0

After the operations in this section are complete, run the **TMOUT=***Timeout interval* command to restore the timeout interval in a timely manner. For example, **TMOUT=600** indicates that a user is logged out if the user does not perform any operation within 600 seconds.

Step 3 Run the following command to generate a key for the node:

- If the node is a Manager management node, run the following command:

  sh ${CONTROLLER_HOME}/sbin/update-ssh-key.sh

- If the node is a non-Manager management node, run the following command:

  sh ${NODE_AGENT_HOME}/bin/update-ssh-key.sh

If "Succeed to update ssh private key." is displayed when the preceding command is executed, the SSH key is generated successfully.

Step 4 Run the following command to copy the public key of the node to the active management node:

scp ${HOME}/.ssh/id_rsa.pub** *oms_ip*:**${HOME}/.ssh/id_rsa.pub_bak

oms_ip: indicates the IP address of the active management node.

Enter the password of user **omm** to copy the files.

Step 5 Log in to the active management node as user **omm**.

Step 6 Run the following command to disable logout on system timeout:

TMOUT=0

Step 7 Run the following command to go to the related directory:

cd ${HOME}/.ssh

Step 8 Run the following command to add new public keys:

cat id_rsa.pub_bak >> authorized_keys

Step 9 Run the following command to move the temporary public key file, for example, **/tmp**.

mv -f id_rsa.pub_bak /tmp

Step 10 Copy the **authorized_keys** file of the active management node to the other nodes in the cluster:

scp authorized_keys *node_ip*:**/${HOME}/.ssh/authorized_keys**

node_ip: indicates the IP address of another node in the cluster. Multiple IP addresses are not supported.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 Run the following command to confirm private key replacement without entering the password:

ssh *node_ip*

node_ip: indicates the IP address of another node in the cluster. Multiple IP addresses are not supported.

Step 12 Log in to FusionInsight Manager. On the homepage, click > **Start** next to the name of the desired cluster to start the cluster. (For MRS 3.3.0 or later, choose **More** > **Start** in the upper right corner of the homepage.)

----End

9.9.2.10 Enabling and Disabling Permission Verification on MRS Cluster Components

Scenario

HDFS and ZooKeeper verify the permission of users who attempt to access the services in both security and normal clusters by default. Users without related permission cannot access resources in HDFS and ZooKeeper. When the cluster is deployed in normal mode, YARN does not verify the permission of users who attempt to access the services by default. All users can access YARN resources.

Based on actual service requirements, administrators can enable YARN permission verification or disable permission verification on HDFS and ZooKeeper in normal clusters.

Notes and Constraints

This topic is available for MRS 3.x or later.

Impact on the System

After the enabling and disabling operations, the service configuration will expire. You need to restart the corresponding service for the configuration to take effect.

Disabling Permission Verification on HDFS

Step 1 Log in to FusionInsight Manager.

Step 2 Click **Cluster**, click the name of the desired cluster, choose **Services** > **HDFS**, and click **Configurations**.

Step 3 Click **All Configurations**.

Step 4 Search for parameters **dfs.namenode.acls.enabled** and **dfs.permissions.enabled**.

- **dfs.namenode.acls.enabled** indicates whether to enable HDFS ACL. The default value is **true**, indicating that the ACL is enabled. Change the value to **false**.
- **dfs.permissions.enabled** indicates whether to enable permission check for HDFS. The default value is **true**, indicating that permission check is enabled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Change the value to **false**. After the modification, the owner, owner group, and permission of the directories and files in HDFS remain unchanged.

Step 5 Click **Save**, click **OK**, and wait for message "Operation successful" to display.

----End

Enabling Permission Verification on YARN

Step 1 Log in to FusionInsight Manager.

Step 2 Click **Cluster**, click the name of the desired cluster, choose **Services** > **Yarn**, and click **Configurations**.

Step 3 Click **All Configurations**.

Step 4 Search for parameter **yarn.acl.enable**.

yarn.acl.enable indicates whether to enable the permission check for YARN.

- In normal clusters, the value is set to **false** by default to disable permission check. To enable permission check, change the value to **true**.
- In security clusters, the value is set to **true** by default to enable authentication.

Step 5 Click **Save**, click **OK**, and wait for message "Operation successful" to display.

----End

Disabling Permission Verification on ZooKeeper

Step 1 Log in to FusionInsight Manager.

Step 2 Click **Cluster**, click the name of the desired cluster, choose **Services** > **ZooKeeper**, and click **Configurations**.

Step 3 Click **All Configurations**.

Step 4 Search for parameter **skipACL**.

skipACL indicates whether to skip the ZooKeeper permission check. The default value is **no**, indicating that permission check is enabled. Change the value to **yes**.

Step 5 Click **Save**, click **OK**, and wait for message "Operation successful" to display.

----End

9.9.2.11 Allowing External Users to Access MRS Clusters in Normal Mode

Scenario

When the cluster is installed in normal mode, the component clients do not support security authentication and cannot use the **kinit** command. Therefore, nodes outside the cluster cannot use users in the cluster by default. This may result in a user authentication failure when one of these nodes access a component server.

The node administrator can configure a user who has the same name as that of a user for a node outside the cluster, allow the user to log in to the node using the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

SSH protocol, and connect to the servers of components in the cluster by using the user who logs in to the OS.

Notes and Constraints

This topic is available for MRS 3.x or later.

Prerequisites

- Nodes outside the cluster can connect to the service plane of the cluster.
- The KrbServer service of the cluster is running properly.
- You have obtained the password of user **root** of the node outside the cluster.
- A human-machine user has been planned and added to the cluster, and you have obtained the authentication credential file. For details, see **Creating a User (MRS 3.x and Later)** and **Downloading MRS Cluster User Credentials**.

Allowing External Users to Access Clusters in Normal Mode

Step 1 Log in to the node where a user is to be added as user **root**.

Step 2 Run the following command:

rpm -qa | grep pam

rpm -qa| grep krb5-client

The following RPM packages are displayed:

```
pam_krb5-32bit-2.3.1-47.12.1
pam-modules-32bit-11-1.22.1
yast2-pam-2.17.3-0.5.211
pam-32bit-1.1.5-0.10.17
pam_mount-32bit-0.47-13.16.1
pam-config-0.79-2.5.58
pam_krb5-2.3.1-47.12.1
pam-doc-1.1.5-0.10.17
pam-modules-11-1.22.1
pam_mount-0.47-13.16.1
pam_ldap-184-147.20
pam-1.1.5-0.10.17
krb5-client-1.6.3
```

Step 3 Check whether the RPM packages in the list are installed in the OS.

- If yes, go to **Step 5**.
- If no, go to **Step 4**.

Step 4 Obtain the lacked RPM packages from the OS image, upload the files to the current directory, and run the following command to install the RPM package:

rpm -ivh *.rpm

WARNING

The RPM packages to be installed may bring security risks. The risks that may be brought by the installation of these RPM packages must be taken into consideration during OS hardening.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the RPM packages are installed, go to **Step 5**.

Step 5 Run the following command to configure Kerberos authentication on PAM:

pam-config --add --krb5

If you need to cancel Kerberos authentication and system user login on a non-cluster node, run the **pam-config --delete --krb5** command as user **root**.

Step 6 Decompress the authentication credential file to obtain **krb5.conf**, use WinSCP to upload this configuration file to the **/etc** directory on the node outside the cluster, and run the following command to configure related permission to enable other users to access the file, such as permission **604**:

chmod 604 /etc/krb5.conf

Step 7 Run the following command in the connection session as user **root** to add the corresponding OS user to the human-machine user, and specify **root** as the primary group.

The OS user password is the same as the initial password when the human-machine user is created on Manager.

useradd *User name* **-m -d /home/admin_test -g root -s /bin/bash**

For example, if the name of the human-machine user is **admin_test**, run the following command:

useradd admin_test -m -d /home/admin_test -g root -s /bin/bash

When you use the newly added OS user to log in to the node by using the SSH protocol for the first time, the system prompts that the password has expired after you enter the user password, and the system prompts that the password needs to be changed after you enter the user password again. You need to enter a new password that meets the password complexity requirements of both the node OS and the cluster.

----End

9.9.2.12 Modifying the Session Timeout Duration for MRS Cluster Management Pages

FusionInsight Manager allows you to configure the timeout duration of the Manager pages. You can adjust the timeout duration according to your service requirements.

You need to set a proper web session timeout duration to avoid potential information leakage resulting from long- term web page exposure.

Notes and Constraints

This function is supported only by MRS 3.3.0 or later.

Changing the Session Timeout Duration

Step 1 Log in to FusionInsight Manager of the MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **OMS**.

Step 3 In the list, locate the row that contains **tomcat** and click **Modify Configuration**.

Step 4 On the displayed page, set **Session Timeout** as required and click **OK**. Enter the user password and wait until the configuration to complete. Once done, the timeout setting will take effect for MRS Manager.

CAUTION

Set the minimum session duration based on service requirements. Otherwise, there will be security risks.

----End

9.9.2.13 Configuring Secure Communication Authorization for an MRS Cluster

MRS clusters provision, manage, and use big data components through the management console. Big data components are deployed in a user's VPC. If the MRS management console needs to directly access big data components deployed in the user's VPC, you need to enable the corresponding security group rules after you have obtained user authorization. This authorization process is called secure communications.

If the secure communications function is not enabled, MRS clusters cannot be created. If you disable the communication after a cluster is created, the cluster status will be **Network channel is not authorized** and the following functions will be affected:

- Installation of cluster components, cluster scaling out or in, upgrading Master node specifications will not be available.
- The cluster running status, alarms, and events cannot be monitored.
- The node management, component management, alarm management, file management, job management, patch management, and tenant management functions on the cluster details page are unavailable.
- The Manager page and the website of each component cannot be accessed.

After the secure communications function is enabled again, the cluster status is restored to **Running**, and the preceding functions become available. For details, see **Enabling Secure Communications for Clusters with This Function Disabled**.

If the security group rules authorized in the cluster are insufficient for you to provision, manage, and use big data components, is displayed on the right of **Secure Communications**. In this case, click **Update** to update the security group rules. For details, see **Updating Secure Communications**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Enabling Secure Communications During Cluster Creation

Step 1 Log in to the MRS console.

Step 2 Click **Create Cluster**. The page for creating a cluster is displayed.

Step 3 On the displayed page, select **Quick Config**.

Step 4 Configure cluster information by referring to **Quickly Creating an MRS Cluster** or **Creating a Custom MRS Cluster**.

Step 5 Select the check box for **Secure Communications**.

Step 6 Click **Create Now**.

If Kerberos authentication is enabled for a cluster, check whether Kerberos authentication is required. If yes, click **Continue**. If no, click **Back** to disable Kerberos authentication and then create a cluster.

----End

Disabling Secure Communications After a Cluster Is Created

Step 1 Log in to the MRS console.

Step 2 In the active cluster list, click the name of the cluster for which you want to disable secure communications.

The cluster details page is displayed.

Step 3 Click the switch on the right of **Secure Communications** to disable authorization. In the dialog box that is displayed, click **OK**.

After the authorization is disabled, the cluster status changes to **Network channel unauthorized**, and some functions of the cluster are unavailable. Exercise caution when performing this operation.

----End

Enabling Secure Communications for Clusters with This Function Disabled

Step 1 Log in to the MRS console.

Step 2 In the active cluster list, click the name of the cluster for which you want to enable secure communications.

The cluster details page is displayed.

Step 3 Click the switch on the right of **Secure Communications** to enable the function.

After the function is enabled, the cluster status changes to **Running**.

----End

Updating Secure Communications

If the security group rules configured for the cluster are insufficient for provisioning, managing, or using big data components, appears to the right of

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Secure Communications. In this case, click **Update** to update the security group rules.

Step 1 Log in to the MRS console.

Step 2 In the active cluster list, click the name of the cluster for which you want to update secure communications.

The cluster details page is displayed.

Step 3 Click **Update** on the right of **Secure Communications**.

Step 4 Click **OK**.

----End

9.9.3 Changing the Passwords for System Users of an MRS Cluster

9.9.3.1 Changing or Resetting the Password for User admin of an MRS Cluster

User **admin** is the system administrator account of Manager. You are advised to periodically change the password by referring to **Changing the Password of User admin** to improve system security. If you forget the password, you can reset it by referring to **Resetting the Password of User admin**.

If the password is changed, the downloaded user credential will be unavailable. Download the authentication credential again, and replace the old one.

Changing the Password of User admin

Changing the password of user admin on Manager:

You can change the password for the user **admin** on Manager for clusters with Kerberos authentication enabled or clusters with Kerberos authentication disabled but with the EIP function enabled.

Step 1 Log in to Manager as user **admin**.

- For MRS 2.*x* or earlier, click the username in the upper right corner of the page and choose **Change Password**.
- For MRS 3.*x* or later, hover over **Hello, admin** in the upper right corner of the page and choose **Change Password**.

Step 2 On the **Change Password** page, set **Old Password**, **New Password**, and **Confirm Password**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The default password complexity requirements are as follows:

- For MRS 2.*x* or earlier:
  - The password must contain at least eight characters.
  - The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
  - The password cannot be the username or the username spelled backwards.
- MRS 3.*x* or later:
  - The password must contain at least eight characters.
  - The password must contain at least four types of the following: uppercase letters, lowercase letters, numbers, spaces, and special characters (~`!?,.;-_'(){}[]/<>@#$%^&*+|\=).
  - The password cannot be the same as the username or the username spelled backwards.
  - The password cannot be a common or easily cracked password.
  - The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**.

Step 3 Click **OK**. Log in to Manager with the new password.

----End

Changing the password of user admin on the cluster node (versions later than MRS 2.*x*):

Step 1 Update the client of the active management node. For details, see **Updating the MRS Cluster Client After the Server Configuration Expires**.

Step 2 Log in to the active management node.

Step 3 (Optional) To change the password as user **omm**, run the following command to switch the user:

sudo su - omm

Step 4 Run the following command to switch to the client directory, for example, **/opt/client**.

cd /opt/client

Step 5 Run the following command to configure environment variables:

source bigdata_env

Step 6 Run the following command to change the password of user **admin**: This operation takes effect in the whole cluster.

kpasswd admin

Enter the old password and then enter a new password twice.

For the cluster, the default password complexity requirements are as follows:

- The password must contain at least eight characters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
- The password cannot be the username or the username spelled backwards.

----End

Resetting the Password of User admin

Step 1 Log in to the **Master1** node.

Step 2 (Optional) To change the password as user **omm**, run the following command to switch the user:

sudo su - omm

Step 3 Go to the client directory.

cd *Client installation directory*

Step 4 Run the following command to set environment variables:

source bigdata_env

Step 5 Run the following command to log in to the console as user **kadmin/admin**:

kadmin -p kadmin/admin

The default password of user **kadmin/admin** is **KAdmin@123** for MRS 2.*x* or earlier and **Admin@123** for MRS 3.*x* or later. The system displays a message indicating that the password has expired upon the first login. Change the password as prompted. Keep the password secure as it cannot be retrieved once lost.

Step 6 Run the following command to reset the password of user **admin**:

cpw admin

For the cluster, the default password complexity requirements are as follows:

- For MRS 2.*x* or earlier:
  - The password must contain at least eight characters.
  - The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
  - The password cannot be the username or the username spelled backwards.
- MRS 3.*x* or later:
  - The password must contain at least eight characters.
  - The password must contain at least four types of the following: uppercase letters, lowercase letters, numbers, spaces, and special characters (~`!?,.;-_'(){}[]/<>@#$%^&*+|\=).
  - The password cannot be the same as the username or the username spelled backwards.
  - The password cannot be a common or easily cracked password.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**.

----End

9.9.3.2 Changing the Passwords for OS Users of an MRS Cluster Node

This section describes how to periodically change the login passwords for the OS users **omm**, **ommdba**, and **root** on MRS cluster nodes to improve system O&M security. You do not need to set a unified password for the OS users on each node.

By default, the password validity period of an OS user is 90 days. **Changing the MRS OS User Password Validity Period** describes how to change the validity period.

You are advised to periodically change a user's login password of the cluster node operating system to improve system O&M security.

Prerequisites

- You have obtained the IP address of the node where the passwords of users **omm** and **ommdba** are to be changed.
- You have obtained the password of user **root** before changing the passwords of users **omm** and **ommdba**.

Changing the Password for OS Users of an MRS Cluster Node

Step 1 Log in to the node that requires password change as user **root**.

Step 2 Run the following command to switch the user:

sudo su - root

Step 3 Run the following command to change the passwords of users **omm**, **ommdba**, or **root**:

passwd omm

passwd ommdba

passwd root

For example, after you run the command to change the password of user **omm**, the following information is displayed:

```
Changing password for user omm.
New password:
Retype new password:
```

Enter the new password and confirm the password. The OS password change policies vary depending on the OS in use.

The default password complexity requirements of an MRS cluster are as follows:

- The password must contain at least eight characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The new password cannot be the same as any of the last five passwords.

----End

Changing the MRS OS User Password Validity Period

Step 1 Log in to the node where you want to change the password validity period of the OS user password as the **root** user.

Step 2 Change the OS user password validity period.

- **Legacy users**

  Run the following command to change the password validity period:

  chage -M *Validity period (days) user_name*

  - *Validity period (days)*: how many days the password is valid since its creation. If this parameter is set to **99999**, the password never expires. Set this parameter based on your service requirements.
  - *user_name*: OS user whose validity period you want to change, for example, **ommdba**.
  - You are advised to set this parameter based on service requirements and periodically change the user password.

- **New users**

  Run the following command to edit the file and change the value of **PASS_MAX_DAYS**, which indicates the password validity period, in days. If the value is changed to **99999**, the password never expires.

  vi /etc/login.defs

----End

9.9.3.3 Changing the Password for the Kerberos Administrator of an MRS Cluster

This section describes how to periodically change the password for the Kerberos or OMS Kerberos administrator **kadmin** of an MRS cluster (MRS 3.*x* or later) to improve system O&M security.

If the password is changed, the downloaded user credential will be unavailable. Download the authentication credential again, and replace the old one.

Prerequisites

- For MRS 2.*x* or earlier, the client has been installed on the **Master1** node.
- For MRS 3.*x* or later, the client has been installed on any node in the cluster and the IP address of the node has been obtained.

Changing the Password of the Kerberos Administrator

If the current MRS version is 3.*x* or later, changing the password of this user will change the password of the OMS Kerberos administrator.

Step 1 Log in to a cluster node.

- For MRS 3.*x* or later, log in to the node where the client is installed as user **root** using the node IP address.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 2.*x* or earlier: Log in to the **Master1** node.

Step 2 (Optional) To change the password as user **omm**, run the following command to switch to user **omm**:

sudo su - omm

Step 3 Run the following command to go to the client directory, for example, **/opt/hadoopclient**:

cd /opt/hadoopclient

Step 4 Run the following command to set environment variables:

source bigdata_env

Step 5 Run the following command to change the password of **kadmin/admin**. This operation takes effect for all servers. Keep the password secure as it cannot be retrieved once lost.

kpasswd kadmin/admin

Enter the password (default: **Admin@123**) and set a new password. The new password must meet the following complexity requirements:

- For MRS 2.*x* or earlier:
  - The password must contain at least eight characters.
  - The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
  - The password cannot be the username or the username spelled backwards.
- MRS 3.*x* or later:
  - The password must contain at least eight characters.
  - The password must contain at least four of the following character types: uppercase letters, lowercase letters, numbers, spaces, and special characters. Supported special characters: ~`!?,.;-_'(){}[]/<>@#$%^&*+|\=
  - The password cannot be the same as the username or the username spelled backwards.
  - The password cannot be a common or easily cracked password.
  - The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**.

----End

Changing the Password of the OMS Kerberos Administrator

This operation is supported only in MRS 3.*x* or later.

Changing this user's password will also update the Kerberos administrator password.

Step 1 Log in to any management node in the cluster as user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Run the following command to go to the directory:

cd ${BIGDATA_HOME}/om-server/om/meta-0.0.1-SNAPSHOT/kerberos/scripts

Step 3 Run the following command to set environment variables:

source component_env

Step 4 Run the following command to change the password of **kadmin/admin**. This operation takes effect for all servers. Keep the password secure as it cannot be retrieved once lost.

kpasswd kadmin/admin

Enter the user password and set a new password. The new password must meet the following complexity requirements:

- The password contains at least 8 characters.
- The password contains at least four types of the following: uppercase letters, lowercase letters, numbers, and special characters (~`!?,.;-_'(){}[]/<>@#$%^&*+|\=).
- The password cannot be the same as the username or the username spelled backwards.
- The password cannot be a common easily-cracked password.
- The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**.

----End

9.9.3.4 Changing the Passwords for Manager Users of an MRS Cluster

Passwords of human-machine system users must be regularly changed to ensure MRS cluster security. This section describes how to change passwords on Manager.

If you have the permission to use Manager, you can change the passwords on Manager.

If you do not have the permission to use Manager, you can change the passwords on the cluster client.

Impact on the System

If you have downloaded a user authentication file, download it again and obtain the keytab file after changing the password of the MRS cluster user.

Prerequisites

- You have obtained the current password policy. For details, see **Configuring Password Policies for MRS Cluster Users**.
- If you need to use a client to reset passwords, ensure that the cluster client has been installed on any node in the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Changing Passwords on Manager

Step 1 Log in to Manager as a user who has the user management permission, for example, **admin**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Hover over the username in the upper right corner of the page.

Choose **Change Password**.

Step 3 On the displayed page, set **Current Password**, **New Password**, and **Confirm Password**, and click **OK**.

The default password complexity requirements are as follows:

- For MRS 2.*x* or earlier:
  - The password must contain at least eight characters.
  - The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
  - The password cannot be the username or the username spelled backwards.
- MRS 3.*x* or later:
  - The password must contain at least eight characters.
  - The password must contain at least four of the following character types: uppercase letters, lowercase letters, numbers, spaces, and special characters. Supported special characters: ~`!?,.;-_'(){}[]/<>@#$%^&*+|\=
  - The password cannot be the same as the username or the username spelled backwards.
  - The password cannot be a common or easily cracked password.
  - The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**.

----End

Changing Passwords on the Client

This function is only available in MRS 3.*x* or later.

Step 1 Log in to the node where the client is installed as the client installation user.

Step 2 Run the following command to switch to the client directory, for example, **/opt/client**:

cd /opt/client

Step 3 Run the following command to set environment variables:

source bigdata_env

Step 4 Change the user password.

kpasswd *System username*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, change the password of the system user **test1**.

kpasswd test1

Enter the old password as prompted and set a new password. The default password complexity requirements are as follows:

- The password must contain at least eight characters.
- The password must contain at least four of the following character types: uppercase letters, lowercase letters, numbers, spaces, and special characters. Supported special characters: ~`!?,.;-_'(){}[]/<>@#$%^&*+|\=
- The password cannot be the same as the username or the username spelled backwards.
- The password cannot be a common or easily cracked password.
- The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**.

If an error occurs during the **kpasswd** command execution, try the following operations:

- Close the SSH session and reopen it.
- Run **kdestroy** and then **kpasswd**.

----End

9.9.3.5 Changing the Password for a Regular LDAP User of an MRS Cluster

This section describes how to periodically change the passwords for LDAP administrators and users of an MRS cluster to improve system O&M security.

- Usernames in MRS 3.1.0:
  - LDAP administrators: cn=root, dc=hadoop, dc=com
  - LDAP users: cn=pg_search_dn, ou=Users, dc=hadoop, dc=com
- Usernames in MRS 2.*x* or earlier:
  - LDAP administrators: rootdn:cn=root, dc=hadoop, dc=com
  - LDAP users: pg_search_dn:cn=pg_search_dn, ou=Users, dc=hadoop, dc=com

Notes and Constraints

- This section applies only to MRS 2.*x* or earlier and MRS 3.1.0. For versions later than MRS 3.1.0, refer to **Modifying the OMS Service Configuration**.
- For MRS 3.1.0 clusters:
  - Changing these user passwords will also update the corresponding OMS LDAP administrator or user passwords.
  - If the cluster is upgraded from an earlier version, the LDAP administrator passwords will retain the password policy of the original cluster. To ensure system security, change the passwords after the upgrade.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- For MRS 2.*x* or earlier, all services need to be restarted after passwords are changed, during which the services are unavailable.
- For MRS 3.1.0 clusters:
  - Changing the user password of the LdapServer service is a high-risk operation and requires restarting the KrbServer and LdapServer services. If KrbServer is restarted, users may fail to be queried by running the **id** command on nodes in the cluster temporarily. Therefore, exercise caution when restarting KrbServer.
  - After the passwords for LDAP users **cn=pg_search_dn**, **ou=Users**, **dc=hadoop**, and **dc=com** are changed, the users may be locked in the LDAP component. You are advised to unlock them after password change. For details, see **Unlocking the LDAP Management Account of the MRS Cluster**.

Prerequisites (MRS 3.1.0)

Before changing the passwords for LDAP users **cn=pg_search_dn**, **ou=Users**, **dc=hadoop**, and **dc=com**, ensure that the users are not locked by running the following command on the active management node of the cluster:

ldapsearch -H ldaps://*Floating IP address of OMS:OLDAP port***-LLL -x -D cn=pg_search_dn,ou=Users,dc=hadoop,dc=com -W -b cn=pg_search_dn,ou=Users,dc=hadoop,dc=com -e ppolicy**

Enter the password for the LDAP user **pg_search_dn**. If the following information is displayed, the user is locked. In this case, unlock the user by referring to **Unlocking the LDAP Management Account of the MRS Cluster**.

```
ldap_bind: Invalid credentials (49); Account locked
```

- To obtain the OLDAP port, log in to FusionInsight Manager, choose **System** > **OMS** > **oldap** > **Modify Configuration**, and view the value of **LDAP Listening Port**.
- The password for the LDAP user **pg_search_dn** is randomly generated by the system, which can be obtained from the **/etc/sssd/sssd.conf** or **/etc/ldap.conf** file on the active node.

Changing the Password for a Regular LDAP User of an MRS Cluster

For MRS 3.1.0:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > **LdapServer**.

Step 3 In the upper right corner of the **Overview** page, choose **More** > **Change Database Password**. In the displayed dialog box, enter the password of the current login user and click **OK**.

Step 4 In the **Change Password** dialog box, select the user whose password needs to be changed in the **User Information** drop-down list.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Enter the old password in the **Old Password** text box, and enter the new password in the **New Password** and **Confirm Password** text boxes.

The default password complexity requirements are as follows:

- The password must contain 16 to 32 characters.
- The password contains at least three types of the following: uppercase letters, lowercase letters, numbers, and special characters (`~!@#$%^&*()-_=+|[{}];,<.>/?).
- The password cannot be the same as the username or the username spelled backwards.
- The password cannot be the same as the current password.

Step 6 Select **I have read the information and understood the impact** and click **OK** to confirm the modification and restart the service.

----End

For MRS 2.*x* or earlier:

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Services** > **LdapServer** > **More**.

Step 3 Click **Change Password**. In the displayed dialog box, enter the old password and click **OK**.

Step 4 In the **Change Password** dialog box, select the user whose password needs to be modified in the **User Information** drop-down box.

Step 5 Enter the old password in the **Old Password** text box, and enter the new password in the **New Password** and **Confirm Password** text boxes.

The default password complexity requirements are as follows:

- The password contains 16 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, and special characters (`~!@#$%^&*()-_=+\|[{}];:",<.>/?).
- The password cannot be the username or the reverse username.
- The new password cannot be the same as the current password.

The default password of the LDAP administrator **rootdn:cn=root,dc=hadoop,dc=com** is **LdapChangeMe@123**, and that of the LDAP user **pg_search_dn:cn=pg_search_dn,ou=Users,dc=hadoop,dc=com** is **pg_search_dn@123**. Periodically change the passwords and keep them secure.

Step 6 Select **I have read the information and understand the impact**, and click **OK** to confirm the modification and restart the service.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.9.3.6 Changing the LDAP Administrator Password for an MRS Cluster

It is recommended that the administrator periodically changes the passwords of LDAP administrator accounts **cn=krbkdc,ou=Users,dc=hadoop,dc=com** and **cn=krbadmin,ou=Users,dc=hadoop,dc=com** to improve the system O&M security.

Notes and Constraints

This section applies only to MRS 3.1.0. For versions later than MRS 3.1.0, see **Modifying the OMS Service Configuration**.

Impact on the System

- You need to restart the KrbServer service after changing the password.
- After the password is changed, check whether the LDAP administrator accounts **cn=krbkdc,ou=Users,dc=hadoop,dc=com** and **cn=krbadmin,ou=Users,dc=hadoop,dc=com** are locked, run the following command on the active management node of the cluster to check whether **krbkdc** is locked (the method for user **krbadmin** is similar):

  ldapsearch -H ldaps://*OMS_FLOAT_ IP address:OLdap port* **-LLL -x -D cn=krbkdc,ou=Users,dc=hadoop,dc=com -W -b cn=krbkdc,ou=Users,dc=hadoop,dc=com -e ppolicy**

  Enter the password of the LDAP administrator account **krbkdc**. The default password is **LdapChangeMe@123**. If the following message is displayed, the account is locked. For details about how to unlock the account, see **Unlocking the LDAP Management Account of the MRS Cluster**.

  ```
  ldap_bind: Invalid credentials (49); Account locked
  ```

  OLdap port number obtaining method:

  a. Log in to FusionInsight Manager, choose **System** > **OMS** > **oldap** > **Modify Configuration**:

  b. The **LDAP Listening Port** parameter value is **oldap port**.

Prerequisites

You have obtained the management node IP address.

Changing the Password of the LDAP Administrator

Step 1 Log in to the active management node as user **omm** with the IP address of the active management node.

Step 2 Run the following command to go to the related directory:

cd ${BIGDATA_HOME}/om-server/om/meta-0.0.1-SNAPSHOT/kerberos/scripts

Step 3 Run the following command to change the password of the LDAP administrator account:

./okerberos_modpwd.sh

Enter the old password and then enter a new password twice.

The password must meet the following complexity requirements:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Contains 16 to 32 characters.
- Contains at least three types of the following: uppercase letters, lowercase letters, numbers, spaces, and special characters (`~!@#$%^&*()-_=+|[{}];,<.>/?).
- Cannot be the same as the current password.

If the following information is displayed, the password is changed.

```
Modify kerberos server password successfully.
```

Step 4 Log in to FusionInsight Manager, click **Cluster**, click the name of the desired cluster, and choose **Services** > **KrbServer**. On the displayed page, choose **More** > **Restart Service**.

Enter the password and do not select **Restart upper-layer services**. Click **OK** to restart the KrbServer service.

----End

9.9.3.7 Changing the Passwords for MRS Cluster Component Running Users

You are advised to regularly change the passwords for MRS cluster component running users to enhance system O&M security.

- For MRS 2.*x* or earlier:
  - If the initial password is randomly generated by the system, reset the password.
  - If the password is changed, the downloaded user credential will be unavailable. Download the authentication credential again, and replace the old one.
- In MRS 3.*x* or later, component running users are classified into the following types based on whether their initial passwords are randomly generated by the system:
  - If the initial password of a component running user is randomly generated by the system, the user is of the machine-machine type.
  - If the initial password of a component running user is not randomly generated by the system, the user is of the human-machine type.

Impact on the System

For MRS 3.*x* or later, the initial password for the component running user is randomly generated by the system. After changing the password, the MRS cluster needs to be restarted, which may cause temporary interruption of services during the restart.

Prerequisites

- For MRS 2.*x* or earlier, the client has been installed on the **Master1** node.
- For MRS 3.*x* or later, the client has been installed on any node in the cluster and the IP address of the node has been obtained.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Changing the Passwords for MRS Cluster Component Running Users (MRS 3.*x* or Later)

Step 1 Log in to the node where the client is installed as the client installation user.

Step 2 Go to the client directory.

cd *Client installation directory*

Step 3 Run the following command to set environment variables:

source bigdata_env

Step 4 Run the following command and enter the password for user **kadmin/admin** to log in to the kadmin console:

kadmin -p kadmin/admin

The default password for the user **kadmin/admin** is **KAdmin@123**, which will expire upon your first login. Change the password as prompted. Keep the password secure as it cannot be retrieved once lost.

Step 5 Run the following command to change the password for an internal component running user:

cpw *Internal component username*

Example: **cpw hdfs**

The username **hdfs** is given as an example. Replace it with the actual one.

The default password complexity requirements are as follows:

- The password contains at least 8 characters.
- The password must contain at least four types of the following: uppercase letters, lowercase letters, numbers, spaces, and special characters (~`!?,.;-_'(){}[]/<>@#$%^&*+|\=).
- The password cannot be the same as the username or the username spelled backwards.
- The password cannot be a common or easily cracked password, for example, **Admin@12345**.
- The new password cannot be the same as any of the last *N* passwords. The value of *N* is determined by **Repetition Rule** in **Configuring Password Policies for MRS Cluster Users**. This policy only applies to human-machine accounts.

Run the following command to view user information:

getprinc *Internal system username*

Example: **getprinc hdfs**

Step 6 Determine the type of the user whose password needs to be changed.

- If the user is a machine-machine user, go to **Step 7**.
- If the user is a human-machine user, the password is successfully changed and no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Log in to FusionInsight Manager.

Step 8 On the home page, click or **More** and click **Restart**.

Step 9 In the displayed dialog box, enter the password of the current login user and click **OK**.

Step 10 In the displayed restart confirmation dialog box, click **OK**.

Step 11 Wait until the system displays a message indicating that the restart is successful.

----End

Changing the Passwords for MRS Cluster Component Running Users (MRS 2.*x* or Earlier)

Step 1 Log in to the **Master1** node.

Step 2 (Optional) To change the password as user **omm**, run the following command to switch the user:

sudo su - omm

Step 3 Run the following command to switch to the client directory, for example, **/opt/client**:

cd /opt/client

Step 4 Run the following command to configure environment variables:

source bigdata_env

Step 5 Run the following command to log in to the console as user **kadmin/admin**:

kadmin -p kadmin/admin

The default password for user **kadmin/admin** is **KAdmin@123**, which will expire upon your first login. Change the password as prompted. Keep the password secure as it cannot be retrieved once lost.

Step 6 Run the following command to reset the password of a component running user. This operation takes effect for all servers.

cpw*Component running user name*

For example, to reset the password of user admin, run the **cpw admin** command.

For the cluster, the default password complexity requirements are as follows:

- The password must contain 8 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, spaces, and special characters '~!@#$%^&*()-_=+\|[{}];:'",<.>/?
- The password cannot be the username or the username spelled backwards.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.9.4 Changing the Passwords for Database Users of an MRS Cluster

9.9.4.1 Changing the Password for the OMS Database Administrator

You are advised to regularly change the password for the OMS database administrator to enhance system O&M security.

Step 1 Log in to the active management node as user **root**.

To ensure proper cluster running, you are not allowed to change the password for the user **ommdba** on the standby management node. Change the password on the active management node only.

Step 2 Run the following command to switch the user:

sudo su - omm

Step 3 Run the following command to switch the directory:

cd $OMS_RUN_PATH/tools

Step 4 Change the password for the user **ommdba**.

mod_db_passwd ommdba

Step 5 Enter the old password of user **ommdba** and enter a new password twice.

The password complexity requirements are as follows:

- The password contains 16 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, and special characters ('~!@#$%^&*()-_=+\|[{}];:'",<.>/?).
- The password cannot be the username or the reverse username.
- The password cannot be the same as the last 20 historical passwords.

If the following information is displayed, the password is changed successfully.

```
Congratulations, update [ommdba] password successfully.
```

----End

9.9.4.2 Changing the Password for an OMS Database Access User

This section describes how to regularly change the password for an OMS database access user to enhance system O&M security.

Impact on the System

The OMS service needs to be restarted for the new password to take effect. The service is unavailable during the restart.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Changing the Password for an OMS Database Access User (MRS 3.*x* or Later)

Step 1 Log in to FusionInsight Manager and choose **System** > **OMS** > **gaussDB** > **Change Password**.

Step 2 Locate the row that contains the user **omm** and click **Change Password** in the **Operation** column.

Step 3 In the displayed dialog box, enter the password of the current login user and click **OK**.

Step 4 Enter the old and new passwords as prompted.

The password complexity requirements are as follows:

- The password must contain 8 to 32 characters.
- The password contains at least three types of the following: uppercase letters, lowercase letters, numbers, and special characters (~`!@#$%^&*()-+_=\| [{}];",<.>/?).
- The password cannot be the same as the username or the username spelled backwards.
- The password cannot be the same as the last 20 historical passwords.

Step 5 Click **OK**. Wait until the system displays a message indicating that the operation is successful.

Step 6 Locate the row that contains the user **omm** and click **Restart OMS Service** in the **Operation** column.

Step 7 In the displayed dialog box, enter the password of the current login user and click **OK**.

Step 8 In the displayed confirmation dialog box, click **OK**.

----End

Changing the Password for an OMS Database Access User (MRS 2.*x* or Earlier)

Step 1 Log in to MRS Manager and click **System**.

Step 2 In the **Permission** area, click **Change OMS Database Password**.

Step 3 Locate the row that contains the user **omm** and click **Change Password** in the **Operation** column.

The password complexity requirements are as follows:

- The password must contain 8 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, and special characters ('~!@#$%^&*()-_=+\| [{}];:'",<.>/?).
- The password cannot be the username or the username spelled backwards.
- The password cannot be the same as the last 20 historical passwords.

Step 4 Click **OK**. When **Operation successful** is displayed, click **Finish**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Locate the row that contains the user **omm** and click **Restart OMS Service** in the **Operation** column.

If the password is changed but the OMS database is not restarted, the user **omm** enters the **Waiting to restart** state, and the password cannot be changed again until the OMS database is restarted.

Step 6 In the displayed dialog box, select **I have read the information and understand the impact**. Click **OK**, and restart the OMS service.

----End

9.9.4.3 Changing the Passwords for Database Users of MRS Cluster Components

This section describes how to regularly change the passwords for database users of cluster components to enhance system O&M security.

Impact on the System

The services need to be restarted for the new password to take effect. The services are unavailable during the restart.

Changing the Passwords for Database Users of MRS Cluster Components (MRS 3.1.0)

This function is only available in MRS 3.1.0. For later versions, refer to **Resetting the MRS Component Database User Password**.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services**.

Step 2 Click the name of the service whose database user password is to be reset. On the displayed **Dashboard** page, click **Stop Service**.

In the displayed dialog box, enter the password for the current login user and click **OK**.

Once you have confirmed the impact of stopping the service, wait for it to come to a complete stop.

Step 3 Click the service whose database user password is to be changed, click **More**, and select **Change Database Password**. On the displayed page, enter the password for the current login user and click **OK**.

Step 4 Enter the old and new passwords as prompted.

The password must meet the following complexity requirements:

- The database user password contains 8 to 32 characters.
- The password contains at least three types of the following: uppercase letters, lowercase letters, numbers, and special characters (~`!@#$%^&*()-+_=\| [{}];",<.>/?).
- The password cannot be the same as the username or the username spelled backwards.
- The password cannot be the same as the last 20 historical passwords.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Select **I have read the information and understand the impact** and click **OK**.

Step 6 After the password is changed, choose **More** > **Restart Service**. In the displayed dialog box, enter the password for the current login user, click **OK**, select **Restart the upper-layer services**, and click **OK**.

----End

Changing the Passwords for Database Users of MRS Cluster Components (MRS 2.*x* or Earlier)

Step 1 On MRS Manager, click **Services** and click the name of the database user service to be modified.

Step 2 Determine the component database user whose password is to be changed.

- To change the password of the DBService database user, go to **Step 3**.
- To change the password for the Hive, Hue, or Loader database user, click **Stop Service** and go to **Step 3**.

Step 3 Choose **More** > **Change Password**.

Step 4 Enter the old and new passwords as prompted.

The password complexity requirements are as follows:

- The password of the DBService database user contains 16 to 32 characters. The password of the Loader, Hive, or Hue database user contains 8 to 32 characters.
- The password must contain at least three types of the following: uppercase letters, lowercase letters, digits, and special characters ('~!@#$%^&*()-_=+\|[{}];:'",<.>/?).
- The password cannot be the username or the reverse username.
- The password cannot be the same as the last 20 historical passwords.

Step 5 Click **OK**. The system automatically restarts the corresponding service. When **Operation successful** is displayed, click **Finish**.

----End

9.9.4.4 Resetting the MRS Component Database User Password

MRS cluster components use random default passwords for connecting to the DBService database. For operations security, you need to periodically reset the passwords of component database users.

Notes and Constraints

This section applies only to MRS 3.1.2 or later. For versions earlier than MRS 3.1.2, see **Changing the Passwords for Database Users of MRS Cluster Components**.

Impact on the System

To reset passwords, you need to stop and then restart services, during which services are unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Resetting the Component Database User Password

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services**.

Step 2 Click the name of the service whose database user password is to be reset, for example, **Kafka**, and click **Stop Service** on the **Dashboard** page.

In the displayed dialog box, enter the password of the current login user and click **OK**.

After confirming the impact of stopping the service, wait until the service is stopped.

Step 3 On the **Dashboard** page, choose **More** > **Reset Database Password**.

In the displayed dialog box, enter the password of the current login user and click **OK**.

Select "I have read the information and understand the impact", and click **OK**.

Step 4 After the password is reset, click **Start Service** on the **Dashboard** page.

Step 5 In the displayed dialog box, click **OK** and wait until the service is started.

----End

9.9.4.5 Resetting the Password for User omm in DBService

Scenario

The default password of the DBService database user **omm** in the MRS cluster is randomly generated. You need to periodically reset the password to improve system O&M security.

Notes and Constraints

This topic is available for MRS 3.2.0-LTS.1 and later versions only. Versions earlier than MRS 3.2.0-LTS.1 do not support password resetting.

Resetting the Password

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **DBService**.

Step 2 On the **Dashboard** page, choose **More** > **Reset Database Password**.

In the displayed dialog box, enter the password of the current login user and click **OK**.

Select **I have read the information and understand the impact**, and click **OK**.

Step 3 After the password is reset, click **More** and select **Service Restart Rolling** on the **Dashboard** page.

In the displayed dialog box, enter the password of the current login user and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Confirm the impact of restarting the service, click **OK**, and wait until the service is started.

----End

9.9.4.6 Changing the Password for User compdbuser of the DBService Database

You need to regularly change the password of the OMS database to enhance system security and maintenance.

Notes and Constraints

This section applies only to MRS 3.x or later.

Changing the Password for the OMS Database Administrator

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **Cluster** > **Services** > **DBService**, click **Instances**, and view the IP address of the active DBService node.

Step 3 Log in to the active DBService node as user **root**.

The password of user **compuserdb** cannot be changed on the standby DBService node. Change the password on the active management node only.

Step 4 Switch to the **$DBSERVER_HOME** directory and configure environment variables.

su - omm

cd $DBSERVER_HOME

source .dbservice_profile

Step 5 Change the password of user **compdbuser** as user **omm** of the DBService database.

gsql -U omm -W omm*Password of user omm of the DBService database* **-d postgres -p 20051 -c "alter user compdbuser identified by '***New password***' valid until '***Expiration time***';"**

- To obtain the initial password of user **omm** in the DBService database, see **MRS Cluster User Accounts** or contact the O&M personnel.
- The new password must meet the following complexity requirements:
  - Contains 16 to 32 characters.
  - Contains at least three types of the following: uppercase letters, lowercase letters, digits, and special characters (~`!@#$%^&*()-+_=\|[{}];:",<.>/?).
  - Cannot be the same as the username or the username spelled backwards.
  - Cannot be the same as the last 20 historical passwords.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The expiration time format is *xxxx-xx-xx*, for example, **2020-10-31**.

If the following information is displayed, the modification is successful:

```
ALTER ROLE
```

----End

9.10 Viewing and Configuring MRS Alarm Events

9.10.1 Viewing MRS Cluster Events

Alarms and events are important mechanisms for ensuring the stability, reliability, and performance of MRS clusters.

Events capture status changes and system operations that occur during system running. They are used to audit or trace system behavior, including component instance startup, shutdown, and active/standby switchover in an MRS cluster, and node-level changes such as slow disk isolation.

For details about common events in an MRS cluster, see **Table 9-102**.

Generally, you do not need to manually handle events. However, **for events of the major or higher severity, you need to check whether the related components are running properly and whether related alarms exist**. For details, see **Viewing Alarms of an MRS Cluster**.

Viewing Cluster Events on the Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 On the **Dashboard** page, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.

Step 4 Choose **Alarms** > **Events**. On the displayed page, view event information about the cluster.

In the event list, you can check events generated in the current cluster, including the event name, severity, and generation time.

You can click the arrow before an event name to expand the event details.

Step 5 To export event information, click **Export**, select the file format, and export the event file to a local directory.

----End

Viewing Cluster Events on Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **O&M** > **Alarm** > **Events**. On the displayed **Events** page, you can view information about all events in the cluster, including the event name, ID, severity, generation time, source, object, and location. By default, the latest 10 events are displayed on each page.

- For clusters of MRS 2.x or earlier, choose **Alarms** > **Events**.
- Click **Export All** to export all event details.
- You can click to manually refresh the current page and click to select the columns that will display.
- You can filter events by object or cluster.
- You can click **Advanced Search** to search for events by event ID, name, severity, start time, or end time.

You can click on the left of an event to view detailed event parameters. **Table 9-101** describes the parameters.

Table 9-101 Event parameters

Parameter	Description
Event ID	Event ID.
Event Name	Event name.
Event Severity	Event severity. The options are **Critical**, **Major**, **Minor**, and **Suggestion**.
Generated	Time when an event is generated.
Object	Possible cause of an event.
Serial Number	Number of events generated by the system.
Location	Detailed information for locating the event, which includes the following: • **Source**: cluster for which the event is generated. • **ServiceName**: service for which the event is generated. • **RoleName**: role for which the event is generated. • **HostName**: host for which the event is generated.
Additional Info	Error information.
Event Cause	Possible cause of an event.
Source	Cluster name.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common Events of an MRS Cluster

Table 9-102 MRS Cluster Event List

Event ID	Component	Event Name	Event Severity
12019	Manager	Stop Service	Warning
12021	Manager	Stop RoleInstance	Warning
12023	Manager	Delete Node	Warning
12024	Manager	Restart Service	Warning
12025	Manager	Restart RoleInstance	Warning
12026	Manager	Manager Switchover	Minor
12065	Manager	Restart Process	Minor
12070	Manager	Job Running Succeeded	Warning
12071	Manager	Job Running Failed	Warning
12072	Manager	Job Killed	Warning
12082	Manager	Automatic Disk Isolation Stopped	Major
12083	Manager	Slow Disk Isolated	Major
12084	Manager	Disk Data Balancing Failed	Major
12085	Manager	Disk Recovered	Major
12086	Manager	Restart Agent	Warning
12087	Manager	Cancel Disk Isolation Failed	Major
12088	Manager	Disk Isolation Cancelled	Major
12089	Manager	Disk Isolation Failed	Major
12090	Manager	Disk Node Isolated	Major
12091	Manager	Disk Node Isolation Cancelled	Major
12092	Manager	Disk Node Instance Started	Major
12093	Manager	Disk Node Isolation Failed	Major
12094	Manager	Disk Node Instance Start Failed	Major
12095	Manager	Cancel Disk Node Isolation Failed	Major
12096	Manager	Disk Node Recovered	Major
12097	Manager	Abnormal Connection to OMS Node Network	Major
12152	Manager	Start Periodic Replication	Minor

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Event ID	Component	Event Name	Event Severity
12153	Manager	Periodic Replication Completed	Minor
12154	Manager	Start Streaming Replication	Minor
12155	Manager	Restart Streaming Replication	Minor
12156	Manager	Stop Streaming Replication	Minor
12157	Manager	Skip Periodic Synchronization	Minor
12158	Manager	Host Information Lost	Minor
14005	HDFS	NameNode Switchover	Minor
14028	HDFS	HDFS Disk Balance	Minor
14029	HDFS	Active NameNode In Safe Mode and New FSimage Generated	Minor
17001	Oozie	Oozie Workflow Execution Failed	Major
17002	Oozie	Oozie Scheduled Job Execution Failed	Major
18001	Yarn	ResourceManager Switchover	Minor
18004	Mapreduce	JobHistoryServer Switchover	Minor
18029	Yarn	Jobs Occupied Too Many Storage Resources	Minor
19001	HBase	HMaster Failover	Minor
19027	HBase	RegionServer-level Hotspot Transfer	Major
19028	HBase	Hotspot Region Splitting	Major
19029	HBase	Hotspot Region Isolation	Major
20003	Hue	Hue Failover	Minor
23002	Loader	Loader Failover	Major
24002	Flume	Flume Channel Overflow	Major
25001	LdapServer	LdapServer Failover	Minor
27000	DBService	DBServer Switchover	Minor
38003	Kafka	Topic Data Storage Period Changed	Warning
43014	Spark	Spark Data Skew	Warning
43015	Spark	Spark SQl Ultra-Large Query Results	Warning

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Event ID	Component	Event Name	Event Severity
43016	Spark	Spark SQL Timeout	Warning
43024	Spark	Start JDBCServer	Warning
43025	Spark	Stop JDBCServer	Warning
43026	Spark	ZooKeeper Connection Succeeded	Warning
43027	Spark	Zookeeper Connection Failed	Warning
43601	GraphBase	GraphBase Failover	Minor
45002	HetuEngine	QAS Failover	Minor
45597	IoTDB	Region Replica Supplementation	Warning
45651	Flink	FlinkServer Failover	Minor

Helpful Links

- You can configure event notifications to receive notification messages through different subscription endpoints (such as SMS messages and emails). For details, see **Configuring Notifications for MRS Cluster Alarms and Events**.
- You can view MRS operation logs on the management console. For details, see **Viewing MRS Operation Logs**.

9.10.2 Viewing Alarms of an MRS Cluster

Alarms and events are important mechanisms for ensuring the stability, reliability, and performance of MRS clusters.

An alarm is a system-generated notification that signals an abnormal condition or fault requiring attention. It is triggered by analyzing system events and requires manual intervention from a user or automatic handling by the system. You can view alarms reported by components on the management console or MRS Manager. You can also define alarm thresholds for threshold-related monitoring metrics in the cluster.

For alarms that can be automatically cleared, the system clears them as soon as the predefined conditions are met. If faults have been rectified and the alarms cannot be automatically cleared, you can manually clear the alarms.

You can view up to 100,000 latest alarms (including uncleared, manually cleared, and automatically cleared alarms) on MRS Manager. If the number of cleared alarms exceeds 100,000 and is about to reach 110,000, the system automatically dumps the earliest 10,000 cleared alarms to the dump path.

The alarm dump directory is as follows. The system automatically generates the directory when alarms are dumped for the first time.

- Clusters of versions earlier than MRS 3.*x*: **${BIGDATA_HOME}/OMSV100R001C00x8664/workspace/data** directory on the active management node

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 3.*x* clusters: **${BIGDATA_HOME}/om-server/OMS/workspace/data** directory on the active management node

Viewing and Clearing Alarms on the Management Console

Step 1 Log in to the MRS console.

Step 2 On the **Active Clusters** page, select a running cluster and click its name to switch to the cluster details page.

Step 3 Click **Alarms** and view the alarm information in the alarm list.

- The alarm list page displays the latest 10 alarms by default.
- You can filter all alarms of the same severity. The results include cleared and uncleared alarms.
- Click **Export All**. In the displayed **Export** dialog box, set **Save As** and click **OK**.

Table 9-103 Alarm descriptions

Parameter	Description
Alarm ID	ID of an alarm.
Alarm Name	Name of an alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Severity	Alarm severity. In versions earlier than MRS 3.x, the cluster alarm severity is as follows: • Critical Indicates alarms reporting errors that affect cluster running, such as unavailable cluster services, node faults, data inconsistency between the active and standby GaussDB databases, and abnormal LdapServer data synchronization. You need to check the cluster status based on the alarms and rectify the faults in a timely manner. • Major Indicates alarms reporting errors that affect some cluster functions, including process faults, periodic backup task failures, and abnormal key file permissions. Check the objects for which the alarms are generated based on the alarms and clear the alarms in a timely manner. • Minor Indicates alarms reporting errors that do not affect major functions of the current cluster, including alarms indicating that the certificate file is about to expire, audit logs fail to be dumped, and the license file is about to expire. • Warning Indicates an alarm of the lowest severity. It is used for information display or prompt and indicates that an event occurs in the scenarios when you stop a service, delete a service, stop an instance, delete an instance, delete a node, restart a service, restart an instance, perform an active/standby switchover for MRS Manager, scale in a host, or restore an instance. Additionally, this type of alarms also occurs when an instance is faulty, a job executed successfully, or a job failed to be executed. In MRS 3.x or later, the alarm severity of a cluster is as follows: • Critical Indicates alarms reporting errors that affect cluster running, such as unavailable cluster services, node faults, data inconsistency between the active and standby GaussDB databases, and abnormal LdapServer data synchronization. You need to check the cluster status based on the alarms and rectify the faults in a timely manner. • Major Indicates alarms reporting errors that affect some cluster functions, including process faults, periodic backup task failures, and abnormal key file permissions. Check the objects for which the alarms are generated based on the alarms and clear the alarms in a timely manner. • Minor Indicates alarms reporting errors that do not affect major functions of the current cluster, including alarms indicating

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
	that the certificate file is about to expire, audit logs fail to be dumped, and the license file is about to expire. • Warning Indicates an alarm of the lowest severity. It is used for information display or prompt and indicates that an event occurs in the scenarios when you stop a service, delete a service, stop an instance, delete an instance, delete a node, restart a service, restart an instance, perform an active/standby switchover for MRS Manager, scale in a host, or restore an instance. Additionally, this type of alarms also occurs when an instance is faulty, a job executed successfully, or a job failed to be executed.
Generated	Time when the alarm is generated.
Location	Details about the alarm.
Operation	If the alarm can be manually cleared, click **Clear Alarm**. To view details about an alarm, click **View Help**. (This function is available in MRS 3.*x* or later).

Step 4 Click **Advanced Search**. In the displayed alarm search area, set search criteria and click **Search** to view information about specified alarms. Click **Reset** to clear the search criteria.

The start time and end time are specified in **Time Range**. You can search for alarms generated within the time range.

Handle the alarm by referring to the **Alarm Reference**. If the alarms in some scenarios are generated due to other cloud services that MRS depends on, you need to contact maintenance personnel of the corresponding cloud services.

Step 5 Click **Clear Alarm** if you need to.

If multiple alarms have been handled, you can select one or more alarms to be cleared and click **Clear Alarm** to clear the alarms in batches. A maximum of 300 alarms can be cleared in each batch.

----End

Viewing and Clearing Alarms on MRS Manager

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Alarm** > **Alarms**.

Step 3 View the alarm information reported by each cluster on FusionInsight Manager, including the alarm name, ID, severity, and generation time. By default, the latest 10 alarms are displayed on each page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 You can click ⌄ on the left of an alarm to view detailed alarm parameters. **Table 9-104** describes the parameters.

Table 9-104 Alarm parameters

Parameter	Description
Alarm ID	Alarm ID.
Alarm Name	Alarm name.
Alarm Severity	Alarm severity. The options are **Critical**, **Major**, **Minor**, and **Suggestion**.
Generated	Time when the alarm is generated.
Cleared	Time when an alarm is cleared. If the alarm is not cleared, **--** is displayed.
Source	Cluster name.
Object	Service, process, or module that triggers the alarm.
Auto Clear	Whether the alarm can be automatically cleared after the fault is rectified.
Alarm Status	Current status of the alarm. The options are **Auto**, **Manual**, and **Uncleared**.
Alarm Cause	Possible cause of an alarm.
Serial Number	Number of alarms generated by the system.
Additional Information	Error information. MRS 3.3.0 or later: You can view the monitoring metric values in **Additional Information** if thresholds are set for the metrics to generate alarms.
Location	Detailed information for locating the alarm, which includes the following: • **Source**: cluster for which the alarm is generated. • **ServiceName**: service for which the alarm is generated. • **RoleName**: role for which the alarm is generated. • **HostName**: host for which the alarm is generated.

Step 5 Manage alarms.

- Click **Export All** to export all alarm details.
- After handling multiple alarms, you can select and clear one or more of them in batches by clicking **Clear Alarm**. Each batch can only clear a maximum of 300 alarms.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- You can filter alarms by object or severity.
- You can click **Advanced Search** to search for alarms by alarm ID, name, type, start time, or end time. Click **Search** to filter alarms that meet the search criteria. Click **Advanced Search** again to view the number of search criteria that you have configured.
- You can click **Clear**, **Mask**, or **View Help** to perform corresponding operations on an alarm.
- If there are a large number of alarms, you can click **View by Category** to sort uncleared alarms by alarm ID. After alarms are classified, click the number of uncleared alarms to view alarm details.

----End

9.10.3 Configuring Alarm Thresholds for an MRS Cluster

Manager allows you to configure monitoring metric thresholds to monitor the health of various metrics. If abnormal data occurs and meets the preset conditions, the system will trigger an alarm, which will appear on the alarm page.

Configuring Alarm Thresholds for a Cluster (MRS 3.*x*)

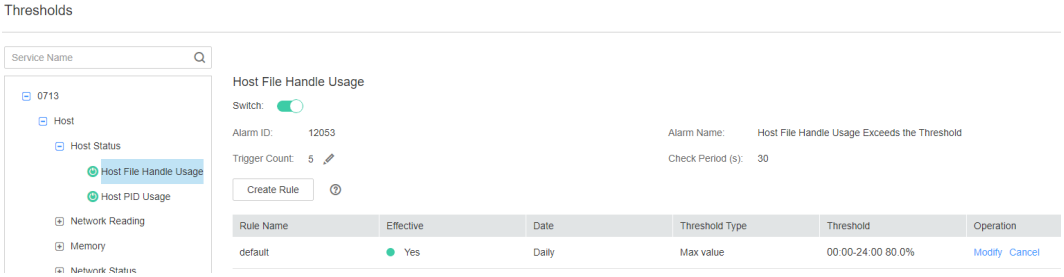
Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Alarm** > **Thresholds**.

Step 3 Select a monitoring metric for a host or service in the cluster.

Figure 9-2 Configuring the threshold for a metric



For example, after selecting **Host Memory Usage**, the information about this indicator threshold is displayed.

- If the alarm sending switch is turned on, an alarm will be triggered if the threshold is reached.
- When **Alarm Severity** is on, hierarchical alarms are enabled. The system dynamically reports alarms at each severity based on the real-time metric values and hierarchical thresholds set for the severity. MRS 3.3.0 or later supports this function.
- **Alarm ID** and **Alarm Name**: alarm information triggered against the threshold
- **Trigger Count**: FusionInsight Manager checks whether the value of a monitoring metric reaches the threshold. If the number of consecutive checks

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

reaches the value of **Trigger Count**, an alarm is generated. **Trigger Count** is configurable.

- **Check Period (s)**: interval for the system to check the monitoring metric.
- The rules in the rule list are used to trigger alarms.

Step 4 Click **Create Rule** to add rules used for monitoring indicators.

Table 9-105 Monitoring indicator rule parameters

Parameter	Description	Example Value
Rule Name	Specifies the alarm rule name.	CPU_MAX

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Alarm Severity	Select an alarm severity. After **Alarm Severity** is on, you need to configure the alarm severity in **Thresholds**. The alarm severity can be one of the following: ● Critical: Indicates a severe fault or issue within the system that could impact cluster running or lead to data loss. This severity level applies to extreme conditions, such as disk or memory usage exceeding 95%. ● Major: Indicates a significant impact on system performance or availability that requires prompt attention to prevent further degradation. ● Minor: Indicates small issues within the system that have no impact on overall system running. Alarms of this severity level should be monitored and addressed, but they have a low handling priority. ● Warning: Indicates the occurrence of a system event and serves purely as an informational notice or prompt.	Major

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Threshold Type	You can use the maximum or minimum value of an indicator as the alarm triggering threshold. If **Threshold Type** is set to **Max value**, the system generates an alarm when the value of the specified indicator is greater than the threshold. If **Threshold Type** is set to **Min value**, the system generates an alarm when the value of the specified indicator is less than the threshold.	Max value
Date	This parameter is used to set the date when the rule takes effect. Possible values are as follows: • Daily • Weekly • Others If **Alarm Severity** is on, only **Daily** is supported.	Daily
Add Date	This parameter is available only when **Date** is set to **Others**. You can set the date when the rule takes effect. Multiple options are available.	09-30
Thresholds	This parameter is used to set the time range when the rule takes effect. If **Alarm Severity** is on, you cannot set the start time and end time. The default start time and end time are 00:00-23:59.	Start and End Time: 00:00–08:30

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
	Thresholds of the rule monitoring metric After **Alarm Severity** is on, different alarm severities can be set for a cluster based on different thresholds. **You can click to set multiple time ranges for the threshold or click to delete one.**	Threshold: 10

Step 5 Click **OK** to save the rules.

Step 6 Locate the row that contains an added rule, and click **Apply** in the **Operation** column. The value of **Effective** for this rule changes to **Yes**.

A new rule can be applied only after you click **Cancel** for an existing rule.

----End

Configuring Alarm Thresholds for a Cluster (Versions Earlier Than MRS 3.*x*)

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Click **System**.

Step 3 In **Configuration**, click **Configure Alarm Threshold** under **Monitoring and Alarm**, select monitoring metrics as planned, and set their baselines.

Step 4 Click a metric, for example, **CPU Usage**, and click **Create Rule**.

Step 5 In the displayed dialog box, set monitoring metric rule parameters.

Table 9-106 Monitoring indicator rule parameters

Parameter	Description	Example Value
Rule Name	Specifies the alarm rule name.	CPU_MAX
Reference Date	Specifies the date on which the reference indicator history is generated.	2014/11/06

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Threshold Type	You can use the maximum or minimum value of an indicator as the alarm triggering threshold. If **Threshold Type** is set to **Max value**, the system generates an alarm when the value of the specified indicator is greater than the threshold. If **Threshold Type** is set to **Min value**, the system generates an alarm when the value of the specified indicator is less than the threshold.	Max value
Alarm Severity	Alarm severity. The value can be: • Critical: Indicates a severe fault or issue within the system that could impact cluster running or lead to data loss. • Major: Indicates a significant impact on system performance or availability that requires prompt attention to prevent further degradation. • Minor: Indicates small issues within the system that have no impact on overall system running. Alarms of this severity level should be monitored and addressed, but they have a low handling priority. • Warning: Indicates the occurrence of a system event and serves purely as an informational notice or prompt.	Major
Time Range	Specifies the period in which the rule takes effect.	From 00:00 to 23:59
Threshold	Specifies the threshold of the rule monitoring metrics.	80
Date	Date when the rule takes effect. • **Workday** • **Weekend** • **Other**	Workday
Add Date	This parameter is valid only when **Date** is set to **Other**. You can select multiple dates.	11/30

Step 6 Click **OK**. A message is displayed in the upper right corner of the page, indicating that the template is saved successfully.

Send alarm is selected by default. **Trigger Count**: FusionInsight Manager checks whether the value of a monitoring metric reaches the threshold. If the number of consecutive checks reaches the value of **Trigger Count**, an alarm is generated. **Trigger Count** is configurable. **Check Period (s)** indicates the interval at which MRS Manager checks monitoring metrics.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Locate the row that contains the newly added rule, and click **Apply** in the **Operation** column. A message is displayed in the upper right corner, indicating that the rule *xx* is successfully applied. Click **Cancel** in the **Operation** column. A message is displayed in the upper right corner, indicating that the rule *xx* is successfully canceled.

----End

Threshold Alarm Metric Reference

FusionInsight Manager alarm monitoring metrics are classified as node information metrics and cluster service metrics. **Table 9-107** lists the metrics whose thresholds can be configured for a node, and **Table 9-108** lists the metrics whose thresholds can be configured for a component.

On FusionInsight Manager of MRS 3.3.0 or later, alarms of some components can be reported by severity. Each alarm severity has a threshold. You can view them on the FusionInsight Manager configuration page.

Table 9-107 Node monitoring metrics

Metric Group	Metric	Description	Default Threshold
CPU	Host CPU Usage	This indicator reflects the computing and control capabilities of the current cluster in a measurement period. By observing the indicator value, you can better understand the overall resource usage of the cluster.	90.0%
Disk	Disk Usage	Indicates the disk usage of a host.	90.0%
	Disk Inode Usage	Indicates the disk inode usage in a measurement period.	80.0%
Memory	Host Memory Usage	Indicates the average memory usage at the current time.	90.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Metric Group	Metric	Description	Default Threshold
Host Status	Host File Handle Usage	Indicates the usage of file handles of the host in a measurement period.	80.0%
	Host PID Usage	Indicates the PID usage of a host.	90%
Network Status	TCP Ephemeral Port Usage	Indicates the usage of temporary TCP ports of the host in a measurement period.	80.0%
Network Reading	Read Packet Error Rate	Indicates the read packet error rate of the network interface on the host in a measurement period.	0.5%
	Read Packet Dropped Rate	Indicates the read packet dropped rate of the network interface on the host in a measurement period.	0.5%
	Read Throughput Rate	Indicates the average read throughput (at MAC layer) of the network interface in a measurement period.	80%
Network Writing	Write Packet Error Rate	Indicates the write packet error rate of the network interface on the host in a measurement period.	0.5%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Metric Group	Metric	Description	Default Threshold
	Write Packet Dropped Rate	Indicates the write packet dropped rate of the network interface on the host in a measurement period.	0.5%
	Write Throughput Rate	Indicates the average write throughput (at MAC layer) of the network interface in a measurement period.	80%
Process	Uninterruptible Sleep Process	Number of D state processes on the host in a measurement period	0
	omm Process Usage	omm process usage in a measurement period	90

Table 9-108 Cluster service indicators

Service	Metric Group	Metric Name	Metric Description	Default Threshold
DBService	Database	Usage of the Number of Database Connections	Usage of database connections	90%
		Disk Space Usage of the Data Directory	Disk space usage of the data directory	80%
Flume	Agent	Heap Memory Usage Calculate	Flume heap memory usage	95.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		Flume Direct Memory Usage Statistics	Flume direct memory usage	80.0%
		Flume Non-heap Memory Usage	Flume non-heap memory usage	80.0%
		Total GC duration of Flume process	Flume total GC time	12000 ms
HBase	GC	GC time for old generation	Total GC time of RegionServer	5000 ms
		GC time for old generation	Total GC time of HMaster	5000 ms
	CPU & memory	RegionServer Direct Memory Usage Statistics	RegionServer direct memory usage	90%
		RegionServer Heap Memory Usage Statistics	RegionServer heap memory usage	90%
		HMaster Direct Memory Usage	HMaster direct memory usage	90%
		HMaster Heap Memory Usage Statistics	HMaster heap memory usage	90%
	Service	Number of Online Regions of a RegionServer	Number of regions of a RegionServer	2000

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		Region in transaction count over threshold	Number of regions that are in the RIT state and reach the threshold duration	1
	Replication	Replication sync failed times (RegionServer)	Number of times that DR data fails to be synchronized	1
		Number of Log Files to Be Synchronized in the Active Cluster	Number of log files to be synchronized in the active cluster	128
		Number of HFiles to Be Synchronized in the Active Cluster	Number of HFiles to be synchronized in the active cluster	128
	Queue	Compaction Queue Size	Size of the Compaction queue	100
HDFS	File and Block	Lost Blocks	Number of block copies that the HDFS lacks of	0
		Blocks Under Replicated	Total number of blocks that need to be replicated by the NameNode	1000
	RPC	Average Time of Active NameNode RPC Processing	Average NameNode RPC processing time	100 ms

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		Average Time of Active NameNode RPC Queuing	Average NameNode RPC queuing time	200 ms
	Disk	HDFS Disk Usage	HDFS disk usage	80%
		DataNode Disk Usage	Disk usage of DataNodes in the HDFS	80%
		Percentage of Reserved Space for Replicas of Unused Space	Percentage of the reserved disk space of all the copies to the total unused disk space of DataNodes.	90%
	Resource	Faulty DataNodes	Indicates the number of faulty DataNodes.	3
		NameNode Non-Heap Memory Usage Statistics	Indicates the percentage of NameNode non-heap memory usage.	90%
		NameNode Direct Memory Usage Statistics	Indicates the percentage of direct memory used by NameNodes.	90%
		NameNode Heap Memory Usage Statistics	Indicates the percentage of NameNode non-heap memory usage.	95%
		DataNode Direct Memory Usage Statistics	Indicates the percentage of direct memory used by DataNodes.	90%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		DataNode Heap Memory Usage Statistics	DataNode heap memory usage	95%
		DataNode Heap Memory Usage Statistics	Indicates the percentage of DataNode non-heap memory usage.	90%
	Garbage Collection	GC Time (NameNode)/ GC Time (DataNode)	Indicates the Garbage collection (GC) duration of NameNodes per minute.	12000 ms
		GC Time	Indicates the GC duration of DataNodes per minute.	12000 ms
Hive	HQL	Percentage of HQL Statements That Are Executed Successfully by Hive	Indicates the percentage of HQL statements that are executed successfully by Hive.	90.0%
	Background	Background Thread Usage	Background thread usage	90%
	GC	Total GC time of MetaStore	Indicates the total GC time of MetaStore.	12000 ms
		Total GC Time in Milliseconds	Indicates the total GC time of HiveServer.	12000 ms
	Capacity	Percentage of HDFS Space Used by Hive to the Available Space	Indicates the percentage of HDFS space used by Hive to the available space.	85.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
	CPU & memory	MetaStore Direct Memory Usage Statistics	MetaStore direct memory usage	95%
		MetaStore Non-Heap Memory Usage Statistics	MetaStore non-heap memory usage	95%
		MetaStore Heap Memory Usage Statistics	MetaStore heap memory usage	95%
		HiveServer Direct Memory Usage Statistics	HiveServer direct memory usage	95%
		HiveServer Non-Heap Memory Usage Statistics	HiveServer non-heap memory usage	95%
		HiveServer Heap Memory Usage Statistics	HiveServer heap memory usage	95%
	Session	Percentage of Sessions Connected to the HiveServer to Maximum Number of Sessions Allowed by the HiveServer	Indicates the percentage of the number of sessions connected to the HiveServer to the maximum number of sessions allowed by the HiveServer.	90.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
Kafka	Partition	Percentage of Partitions That Are Not Completely Synchronized	Indicates the percentage of partitions that are not completely synchronized to total partitions.	50%
	Others	Unavailable Partition Percentage	Percentage of unavailable partitions of each Kafka topic	40%
		User Connection Usage on Broker	Usage of user connections on Broker	80%
	Disk	Broker Disk Usage	Indicates the disk usage of the disk where the Broker data directory is located.	80.0%
		Disk I/O Rate of a Broker	I/O usage of the disk where the Broker data directory is located	80%
	Process	Broker GC Duration per Minute	Indicates the GC duration of the Broker process per minute.	12000 ms
		Heap Memory Usage of Kafka	Indicates the Kafka heap memory usage.	95%
		Kafka Direct Memory Usage	Indicates the Kafka direct memory usage.	95%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
Loader	Memory	Heap Memory Usage Calculate	Indicates the Loader heap memory usage.	95%
		Direct Memory Usage of Loader	Indicates the Loader direct memory usage.	80.0%
		Non-heap Memory Usage of Loader	Indicates the Loader non-heap memory usage.	80%
	GC	Total GC time of Loader	Indicates the total GC time of Loader.	12000 ms
MapReduce	Garbage Collection	GC Time	Indicates the GC time.	12000 ms
	Resource	JobHistoryServer Direct Memory Usage Statistics	Indicates the JobHistoryServer direct memory usage.	90%
		JobHistoryServer Non-Heap Memory Usage Statistics	Indicates the JobHistoryServer non-heap memory usage.	90%
		JobHistoryServer Heap Memory Usage Statistics	Indicates the JobHistoryServer non-heap memory usage.	95%
Oozie	Memory	Oozie Heap Memory Usage Calculate	Indicates the Oozie heap memory usage.	95.0%
		Oozie Direct Memory Usage	Indicates the Oozie direct memory usage.	80.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		Oozie Non-heap Memory Usage	Indicates the Oozie non-heap memory usage.	80%
	GC	Total GC duration of Oozie	Indicates the Oozie total GC time.	12000 ms
Spark/ Spark2x	Memory	JDBCServer2x Heap Memory Usage Statistics	JDBCServer2x heap memory usage	95%
		JDBCServer2x Direct Memory Usage Statistics	JDBCServer2x direct memory usage	95%
		JDBCServer2x Non-Heap Memory Usage Statistics	JDBCServer2x non-heap memory usage	95%
		JobHistory2x Direct Memory Usage Statistics	JobHistory2x direct memory usage	95%
		JobHistory2x Non-Heap Memory Usage Statistics	JobHistory2x non-heap memory usage	95%
		JobHistory2x Heap Memory Usage Statistics	JobHistory2x heap memory usage	95%
		IndexServer2x Direct Memory Usage Statistics	IndexServer2x direct memory usage	95%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		IndexServer2x Heap Memory Usage Statistics	IndexServer2x heap memory usage	95%
		IndexServer2x Non-Heap Memory Usage Statistics	IndexServer2x non-heap memory usage	95%
	GC Count	Full GC Number of JDBCServer2x	Full GC times of JDBCServer2x	12
		Full GC Number of JobHistory2x	Full GC times of JobHistory2x	12
		Full GC Number of IndexServer2x	Full GC times of IndexServer2x	12
	GC Time	Total GC Time in Milliseconds	Total GC time of JDBCServer2x	12000 ms
		Total GC Time in Milliseconds	Total GC time of JobHistory2x	12000 ms
		Total GC Time in Milliseconds	Total GC time of IndexServer2x	12000 ms
Storm	Cluster	Number of Available Supervisors	Indicates the number of available Supervisor processes in the cluster in a measurement period.	1
		Slot Usage	Indicates the slot usage in the cluster in a measurement period.	80.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
	Nimbus	Nimbus Heap Memory Usage Calculate	Indicates the Nimbus heap memory usage.	80%
Yarn	Resources	NodeManager Direct Memory Usage Statistics	Indicates the percentage of direct memory used by NodeManagers.	90%
		NodeManager Heap Memory Usage Statistics	Indicates the percentage of NodeManager heap memory usage.	95%
		NodeManager Non-Heap Memory Usage Statistics	Indicates the percentage of NodeManager non-heap memory usage.	90%
		ResourceManager Direct Memory Usage Statistics	Indicates the ResourceManager direct memory usage.	90%
		ResourceManager Heap Memory Usage Statistics	Indicates the ResourceManager heap memory usage.	95%
		ResourceManager Non-Heap Memory Usage Statistics	Indicates the ResourceManager non-heap memory usage.	90%
	Garbage collection	GC Time	Indicates the GC duration of NodeManager per minute.	12000 ms

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		GC Time	Indicates the GC duration of ResourceManager per minute.	12000 ms
	Others	Failed Applications of root queue	Number of failed tasks in the root queue	50
		Terminated Applications of root queue	Number of killed tasks in the root queue	50
	CPU & memory	Pending Memory	Pending memory capacity	83886080 MB
	Application	Pending Applications	Pending tasks	60
ZooKeeper	Connection	ZooKeeper Connections Usage	Indicates the percentage of the used connections to the total connections of ZooKeeper.	80%
	CPU & memory	Heap Memory Usage	Indicates the ZooKeeper heap memory usage.	95%
		Heap Memory Usage Calculate	Indicates the ZooKeeper direct memory usage.	80%
	GC	ZooKeeper GC Duration per Minute	Indicates the GC time of ZooKeeper every minute.	12000 ms
Ranger	GC	UserSync GC Duration	UserSync garbage collection (GC) duration	12000 ms
		RangerAdmin GC Duration	RangerAdmin GC duration	12000 ms

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
		TagSync GC Duration	TagSync GC duration	12000 ms
	CPU & memory	UserSync Non-Heap Memory Usage	UserSync non-heap memory usage	80.0%
		UserSync Direct Memory Usage	UserSync direct memory usage	80.0%
		UserSync Heap Memory Usage	UserSync heap memory usage	95.0%
		RangerAdmin Non-Heap Memory Usage	RangerAdmin non-heap memory usage	80.0%
		RangerAdmin Heap Memory Usage	RangerAdmin heap memory usage	95.0%
		RangerAdmin Direct Memory Usage	RangerAdmin direct memory usage	80.0%
		TagSync Direct Memory Usage	TagSync direct memory usage	80.0%
		TagSync Non-Heap Memory Usage	TagSync non-heap memory usage	80.0%
		TagSync Heap Memory Usage	TagSync heap memory usage	95.0%

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Metric Group	Metric Name	Metric Description	Default Threshold
ClickHouse	Cluster Quota	Quota of the ZooKeeper Nodes Used by a ClickHouse Service	Quota of the ZooKeeper nodes used by a ClickHouse service.	90%
		Capacity Quota of ZooKeeper Used by a ClickHouse Service	Capacity quota of ZooKeeper directory used by the ClickHouse service	90%
IoTDB	GC	IoTDBServer GC Duration	IoTDBServer garbage collection (GC) duration	12000 ms
	CPU & memory	IoTDBServer Heap Memory Usage	IoTDBServer heap memory usage	90%
		IoTDBServer Direct Memory Usage	IoTDBServer direct memory usage	90%

9.10.4 Configuring Alarm Masking for an MRS Cluster

MRS clusters support manual alarm masking, allowing administrators to filter out redundant notifications and focus on critical alarms more effectively.

- Some unimportant alarms and minor alarms need to be masked during MRS cluster running.
- When a third-party product is interconnected with MRS, some alarms of the product are duplicated with the alarms of MRS and need to be masked.
- When the deployment environment is special, certain alarms may be falsely reported and need to be masked.

Once an alarm is masked, any new alarms with the same ID will no longer appear on the **Alarm Management** page or be counted. However, previously reported alarms will still be displayed.

Configuring the Alarm Masking Status

Step 1 Log in to FusionInsight Manager of the MRS cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

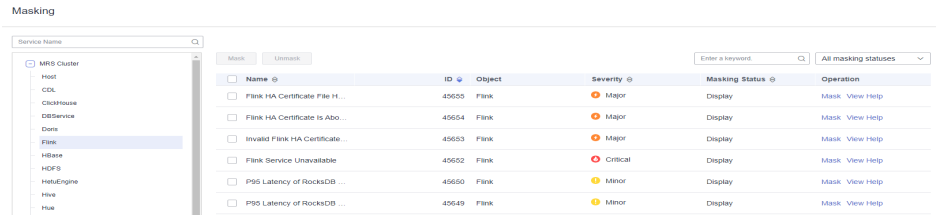
For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **O&M** > **Alarm** > **Masking Setting**.

Step 3 In the **Masking Setting** area, select the specified service or module.

Step 4 Select an alarm from the alarm list.

Figure 9-3 Masking an alarm



The information about the alarm is displayed, including the alarm name, ID, severity, masking status, and operations can be performed on the alarm.

- The masking status includes **Display** and **Masking**.
- Operations include **Masking** and **Help**.

Step 5 Set the masking status for an alarm:

- Click **Masking**. In the displayed dialog box, click **OK** to change the alarm masking status to **Masking**.
- Click **Cancel Masking**. In the dialog box that is displayed, click **OK** to change the masking status of the alarm to **Display**.

----End

Configuring Alarm Masking and Alarm Trigger Count

MRS allows you to mask alarms and number of trigger times on the background. If the number of alarm report times is less than or equal to the number of alarm trigger times, no alarm will be reported.

The following steps apply only to MRS 3.5.0 or later clusters.

Step 1 Log in to the active management node as user **root**.

For details about how to log in to an MRS cluster node, see **Logging In to an MRS Cluster Node**.

Step 2 Run the following command to switch to user **omm**:

```
su - omm
```

Step 3 Run the following command to modify the **alarm_filter_config.json** configuration file. If the file does not exist, create it.

```
vi $BIGDATA_HOME/om-server/OMS/workspace/conf/fms/alarm_filter_config.json
```

Step 4 Configure or add parameters based on service demands.

- **Alarm ID**: ID of the alarm to be masked, for example, **12016**.
- **is_filtered:** Whether to mask the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

true indicates that the alarm is masked and not reported. **false** indicates that the alarm is reported after certain times of trigger.

- **smoothing_times**: trigger count. If the number of alarm trigger times is less than or equal to the value of this parameter, no alarm will be reported. The value is an integer greater than 0.

The **is_filtered** parameter takes priority over the **smoothing_times** parameter.

For example, if the configuration is as follows, alarm 12016 is masked and alarm 12017 is reported after it is triggered for more than three times:

```
{
  "12016": {
    "is_filtered": true,
    "smoothing_times": 1
  },
  "12017": {
    "is_filtered": false,
    "smoothing_times": 3
  }
}
```

Step 5 Restart the FMS service on the active OMS node to apply the changes.

```
sh $BIGDATA_HOME/om-server/OMS/workspace/bin/omm_s_fm_ctl.sh restart
```

Step 6 Log in to the standby management node as user **root** and perform **Step 3** and **Step 4**.

----End

9.10.5 Connecting an MRS Cluster to SNMP to Report Alarms

If users need to view alarms and monitoring data of a cluster on the O&M platform, you can use Simple Network Management Protocol (SNMP) on FusionInsight Manager to report related data to the network management system (NMS).

Prerequisites

The ECS corresponding to the server must be in the same VPC as the MRS cluster's Master node, and the Master node must be able to access the server's IP address and specified port.

Connecting an MRS Cluster to SNMP to Report Alarms

Step 1 Log in to MRS Manager and access the SNMP configuration page.

- MRS 3.*x* or later:
  a. Log in to FusionInsight Manager.
  b. Choose **System** > **Interconnection** > **SNMP**.
  c. Enable the SNMP service.
- Versions earlier than MRS 3.*x*:
  a. Log in to MRS Manager and click **System**.
  b. In **Configuration**, enable **Configure SNMP** under **Monitoring and Alarm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Set interconnection parameters based on **Table 9-109**.

Table 9-109 Interconnection parameters

Parameter	Example Value	Description
Version	V3	Specifies the version of SNMP, which can be: • **V2C**: This is an earlier version with low security. • **V3**: This is a later version with higher security than SNMPv2c SNMPv3 is recommended.
Local Port	20000	Specifies the local port. The default value is **20000**. The value ranges from **1025** to **65535**.
Read Community Name	-	Specifies the read-only community name. This parameter is available only when **Version** is set to **V2C**.
Write Community Name	-	Specifies the write community name. This parameter is available only when **Version** is set to **V2C**.
Security Username	test	Specifies the SNMP security username. This parameter is available only when **Version** is set to **V3**.
Authentication Protocol	SHA	Specifies the authentication protocol. This parameter is available only when **Version** is set to **V3**. SHA is recommended.
Authentication Password	-	Specifies the authentication password. This parameter is available only when **Version** is set to **V3**.
Confirm Password	-	Used to confirm the authentication password. This parameter is available only when **Version** is set to **V3**.
Encryption Protocol	AES256	Specifies the encryption protocol. This parameter is available only when **Version** is set to **V3**. AES256 is recommended.
Encryption Password	-	Specifies the encryption password. This parameter is available only when **Version** is set to **V3**.
Confirm Password	-	Used to confirm the encryption password. This parameter is available only when **Version** is set to **V3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The value of **Security Username** cannot contain repeated strings with the unit length as a common factor of 64 (such as 1, 2, 4, and 8), for example, **abab** and **abcdabcd**.
- The **Authentication Password** and **Encryption Password** must contain 8 to 16 characters, including at least three types of the following characters: uppercase letters, lowercase letters, digits, and special characters. The two passwords must be different. The two passwords cannot be the same as the security username or username spelled backwards.
- For security purposes, periodically change the authentication password and encryption password when the SNMP protocol is used.
- If SNMPv3 is used, a security user will be locked after five consecutive authentication failures within 5 minutes. The user will be automatically unlocked 5 minutes later.

Step 3 Click **Create Trap Target** in the **Trap Target** area. In the displayed dialog box, set the following parameters:

- **Target Symbol**: specifies the trap target ID, which is the ID of the NMS or host that receives traps. The value consists of 1 to 255 characters, including letters or digits.
- **Target IP Address Mode** (only for MRS 3.*x* or later): mode of the target IP address. The options are **IPv4** and **IPv6**.
- **Target IP Address**: specifies the target IP address, which can communicate with the management plane IP address of the management node.
- **Target Port**: specifies the port receiving traps. The port number must be consistent with the peer end and ranges from 0 to 65535.
- **Trap Community Name**: This parameter is available only when **Version** is set to **V2C** and is used to report the community name.

Click **OK**.

The **Create Trap Target** dialog box is closed.

Step 4 Click **OK**.

----End

9.10.6 Connecting an MRS Cluster to the Syslog Server to Report Alarms

If users need to view alarms and events of a cluster on the unified alarm reporting platform, you can use the Syslog protocol to report related data to the alarm platform.

WARNING

The Syslog protocol is not inherently encrypted, making data transmitted via Syslog vulnerable to theft and security breaches. In high-security environments, it is crucial to use encryption and authentication mechanisms to protect sensitive information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

The ECS corresponding to the server must be in the same VPC as the MRS cluster's Master node, and the Master node must be able to access the server's IP address and specified port.

Connecting an MRS Cluster to the Syslog Server to Report Alarms

Step 1 Log in to MRS Manager by referring to **Accessing MRS Manager** and access the SNMP configuration page.

- MRS 3.*x* or later:
  a. Log in to FusionInsight Manager.
  b. Choose **System** > **Interconnection** > **Syslog**.
  c. Enable the Syslog service.
- Versions earlier than MRS 3.*x*:
  a. Log in to MRS Manager and click **System**.
  b. In **Configuration**, enable **Configure Syslog** under **Monitoring and Alarm**.

Step 2 Set northbound parameters based on **Table 9-110**.

Table 9-110 Syslog interconnection parameters

Parameter Area	Parameter	Example Value	Description
Syslog Protocol	Server IP Address Mode	IPV4	Specifies the IP address mode of the interconnected server. The value can be **IPV4** or **IPV6**. (only for MRS 3.*x* or later)
	Server IP Address	192.168.*xx*.*xxx*	Specifies the IP address of the interconnected server.
	Server Port	514	Specifies the port number for interconnection.
	Protocol	UDP	Specifies the protocol type, including: • **TCP** • **UDP**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter Area	Parameter	Example Value	Description
	Severity Level	Informational	Specifies the severity of the reported message. The options are as follows: • **Emergency** • **Alert** • **Critical** • **Error** • **Warning** • **Notice** • **Informational** (default value) • **Debug** **NOTE** **Severity Level** and **Facility** determine the priority of the sent message. **Priority** = **Facility** × 8 + **Severity Level** For the values of **Severity Level** and **Facility**, see **Table 9-111**.
	Facility	local use 0 (local0)	Specifies the module where the log is generated. For the available values of this parameter, see **Table 9-111**. Default value **local use 0 (local0)** is recommended.
	Identifier	FusionInsight Manager	Product ID. The identifier can contain a maximum of 256 characters, including letters, digits, underscores (_), periods (.), hyphens (-), spaces, and the following special characters: | $ { } • MRS 3.*x* or later: The default value is **FusionInsight Manager**. • Versions earlier than MRS 3.*x*: The default value is **MRS Manager**.
Report Message	Report Format	${alarm.id}	Specifies the message format of the alarm report. For details, see the help information on the page. The report format can contain a maximum of 1024 characters, including letters, digits, underscores (_), periods (.), hyphens (-), spaces, and the following special characters: | $ { } **NOTE** For details about each field in the report format, see **Table 9-112**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter Area	Parameter	Example Value	Description
	Alarm Type	Fault	Specifies the type of the alarm to be reported. • For MRS 3.*x* or later, see **Table 9-112**. • Versions earlier than MRS 3.*x*: – **Fault**: indicates that the Syslog alarm message is reported when MRS Manager generates an alarm. – **Clear**: indicates that a Syslog alarm message is reported when an alarm on MRS Manager is cleared. – **Event**: indicates that the Syslog alarm message is reported when MRS Manager generates an event.
	Alarm Severities	Major	Specifies the level of the alarm to be reported. • For MRS 3.*x* or later, see **Table 9-112**. • Versions earlier than MRS 3.*x*: **Suggestion**, **Minor**, **Major**, and **Critical**.
Uncleared Alarm Reporting	Periodic Uncleared Alarm Reporting	Enable	Specifies whether to report uncleared alarms in a specified period. You can toggle on or off the function. The function is toggled off by default.
	Report Interval (min)	30	Specifies the interval for periodically reporting uncleared alarms. This parameter is valid only when **Periodic Uncleared Alarm Reporting** is enabled. The default value is **15**, in minutes. The value ranges from **5** to **1440** (one day).
Heartbeat Settings	Heartbeat Reporting	Enable	Specifies whether to periodically report Syslog heartbeat messages. You can toggle on or off the function. The function is toggled off by default.
	Heartbeat Interval (minutes)	15	Specifies the interval for periodically reporting heartbeat messages. This parameter is valid only when **Heartbeat Reporting** is enabled. The default value is **15**, in minutes. The value ranges from **1** to **60**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter Area	Parameter	Example Value	Description
	Heartbeat Packet	FusionInsight Syslog Heartbeat	Specifies the heartbeat message to be reported. This parameter is valid when **Heartbeat Reporting** is toggled on and cannot be left blank. The value can contain a maximum of 256 characters, including digits, letters, underscores (_), vertical bars (|), colons (:), spaces, commas (,), and periods (.).

After periodic heartbeat packet reporting is enabled, packet sending may be interrupted during some automatic cluster fault recovery scenarios (for example, active/standby OMS switchover). The sending process resumes automatically after the recovery.

Step 3 Click **OK**.

----End

Related Information

Table 9-111 Numeric codes of **Severity Level** and **Facility**

Severity Level	Facility	Numeric Code
Emergency	kernel messages	0
Alert	user-level messages	1
Critical	mail system	2
Error	system daemons	3
Warning	security/authorization messages (note 1)	4
Notice	messages generated internally by syslog	5
Informational	line printer subsystem	6
Debug	network news subsystem	7
-	UUCP subsystem	8
-	clock daemon (note 2)	9
-	security/authorization messages (note 1)	10
-	FTP daemon	11

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Severity Level	Facility	Numeric Code
-	NTP subsystem	12
-	log audit (note 1)	13
-	log alert (note 1)	14
-	clock daemon (note 2)	15
-	local use 0~7 (local0 ~ local7)	16 to 23

Table 9-112 Report format information fields

Information Field	Description
dn	Cluster name
id	Alarm ID
name	Alarm name
serialNo	Alarm serial number The fault alarms and their corresponding clear alarms share the same serial number.
category	Alarm type. The options are as follows: ● **0**: fault alarm ● **1**: clear alarm ● **2**: event
occurTime	Time when the alarm was generated
clearTime	Time when this alarm was cleared
isAutoClear	Whether an alarm is automatically cleared. The options are as follows: ● **1**: yes ● **0**: no
locationInfo	Location where the alarm was generated
clearType	Alarm clearance type. The options are as follows: ● **-1**: not cleared ● **0**: automatically cleared ● **2**: manually cleared

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Information Field	Description
level	Severity. The options are as follows: • **1**: critical alarm • **2**: major alarm • **3**: minor alarm • **4**: warning alarm
cause	Alarm cause
additionalInfo	Additional information
object	Alarm object

9.10.7 Periodically Backing Up Alarm and Audit Information

You can modify the configuration file to periodically back up FusionInsight Manager alarm information, FusionInsight Manager audit information, and audit information of all services to the specified storage location.

The backup can be performed using FTP or SFTP. FTP does not encrypt data, which may cause security risks. Therefore, SFTP is recommended.

Notes and Constraints

This section applies only to MRS 3.*x* or later.

Periodically Backing Up Alarm and Audit Information

Step 1 Log in to the active management node of an MRS cluster as user **root**.

For details about how to log in to an MRS cluster node, see **Logging In to an MRS Cluster**.

Perform this operation only on the active management node. Scheduled backup is not supported on the standby management node.

Step 2 Run the following command to switch to user **omm**:

```
su - omm
```

Step 3 Run the following command to switch the directory:

```
cd ${BIGDATA_HOME}/om-server/om/sbin
```

Step 4 Run the following command to configure scheduled backup of FusionInsight Manager's alarm and audit information or service audit information:

./setNorthBound.sh -t *Information type* **-i** *Remote server IP address* **-p** *SFTP or FTP port used by the server*-**u** *Username* **-d** *Save path* **-c** *Interval (minutes)* **-m** *Number of records in each file* **-s** *Whether to enable backup* **-e** *Protocol*

Example:

```
./setNorthBound.sh -t alarm -i 10.0.0.10 -p 22 -u sftpuser -d /tmp/ -c 10 -m 100 -s true -e sftp
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This script modifies the alarm backup configuration file **alarm_collect_upload.properties**.

The file save path is **${BIGDATA_HOME}/om-server/tomcat/webapps/web/WEB-INF/classes/config**.

```
./setNorthBound.sh -t audit -i 10.0.0.10 -p 22 -u sftpuser -d /tmp/ -c 10 -m 100  -s true -e sftp
```

This script modifies the audit backup configuration file **audit_collect_upload.properties**.

The file save path is **${BIGDATA_HOME}/om-server/tomcat/webapps/web/WEB-INF/classes/config**.

```
./setNorthBound.sh -t service_audit -i 10.0.0.10 -p 22 -u sftpuser -d /tmp/ -c 10 -m 100  -s true -e sftp
```

This script modifies the service audit backup configuration file **service_audit_collect_upload.properties**.

The file save path is **${BIGDATA_HOME}/om-server/tomcat/webapps/web/WEB-INF/classes/config**.

Step 5 Enter the password as prompted.

The password is encrypted and saved in the configuration file.

```
Please input sftp/ftp server password:
```

Step 6 Check the configuration result. If the following information is displayed, the configuration is successful. The configuration file will be automatically synchronized to the standby management node.

```
execute command syncfile successfully.
Config Succeed.
```

----End

9.10.8 Enabling the MRS Cluster Maintenance Mode to Disable Alarm Reporting

FusionInsight Manager allows you to set clusters, services, hosts, or OMSs to the maintenance mode. Objects in maintenance mode do not report alarms. This prevents the system from generating a large number of unnecessary alarms during maintenance changes, such as upgrade, because these alarms may influence O&M personnel's judgment on the cluster status.

- Cluster maintenance mode

  If a cluster is not brought online or has been brought offline due to O&M operations (for example, non-rolling upgrade), you can set the entire cluster to the maintenance mode.

- Service maintenance mode

  When performing maintenance operations on a specific service (for example, performing service-affecting commissioning operations like batch restart of service instances, directly powering on or off nodes of the service, or repairing the service), you can set only this service to the maintenance mode.

- Host maintenance mode

  When performing maintenance operations on a host (such as powering on or off, isolating, or reinstalling the host, upgrading its OS, or replacing the host), you can set only this host to the maintenance mode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- OMS maintenance mode

  When restarting, replacing, or repairing an OMS node, you can set the OMS node to the maintenance mode.

Notes and Constraints

Only MRS 3.*x* and later clusters support the cluster maintenance mode.

Impact on the System

After the maintenance mode is set, alarms caused by non-maintenance operations are suppressed and cannot be reported. Alarms can be reported only when faults persist after the system exits the maintenance mode. Therefore, exercise caution when setting the maintenance mode.

Enabling the MRS Cluster Maintenance Mode

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Set the maintenance mode.

Determine the object to set the maintenance mode based on the service scenario. For details, see **Table 9-113**.

Table 9-113 Setting to the maintenance mode

Scenario	Operation
Configure a cluster to enter the maintenance mode.	1. On the management page, choose ··· or **More** > **Enter Maintenance**. 2. In the displayed dialog box, click **OK**. After the cluster enters the maintenance state, the status of the cluster becomes . After maintenance is complete, click **Exit Maintenance Mode**. The cluster then exits the maintenance mode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Scenario	Operation
Configure a service to enter the maintenance mode.	1. On FusionInsight Manager, choose **Cluster** > **Services** > *Service name*. 2. On the service details page, click **More** and select **Enter Maintenance Mode**. 3. In the displayed dialog box, click **OK**. After a service enters the maintenance mode, the status of the service becomes in the service list. After maintenance is complete, click **Exit Maintenance Mode**. The service then exits the maintenance mode. **When configuring a service to enter the maintenance mode, you are advised to set the upper-layer services that depend on this service to the maintenance mode as well.**
Configure a host to the maintenance mode.	1. On FusionInsight Manager, choose **Hosts**. 2. On the **Hosts** page, select the target host, click **More**, and select **Enter Maintenance Mode**. 3. In the displayed dialog box, click **OK**. After the host enters the maintenance mode, the status of the host becomes in the host list. After maintenance is complete, click **Exit Maintenance Mode**. The host then exits the maintenance mode.
Configure the OMS to enter the maintenance mode.	1. On FusionInsight Manager, choose **System** > **OMS** > **Enter Maintenance Mode**. 2. In the displayed dialog box, click **OK**. After the OMS enters the maintenance state, the OMS status becomes . After maintenance is complete, click **Exit Maintenance Mode**. The OMS then exits the maintenance mode.

Step 3 Check the cluster maintenance view.

On FusionInsight Manager, click or **More** next to the cluster name and select **Maintenance Mode View**. In the displayed window, you can view the services and hosts in maintenance mode in the cluster.

After maintenance is complete, you can select services and hosts in batches in the maintenance mode view and click **Exit Maintenance Mode** to make them exit the maintenance mode.

----End

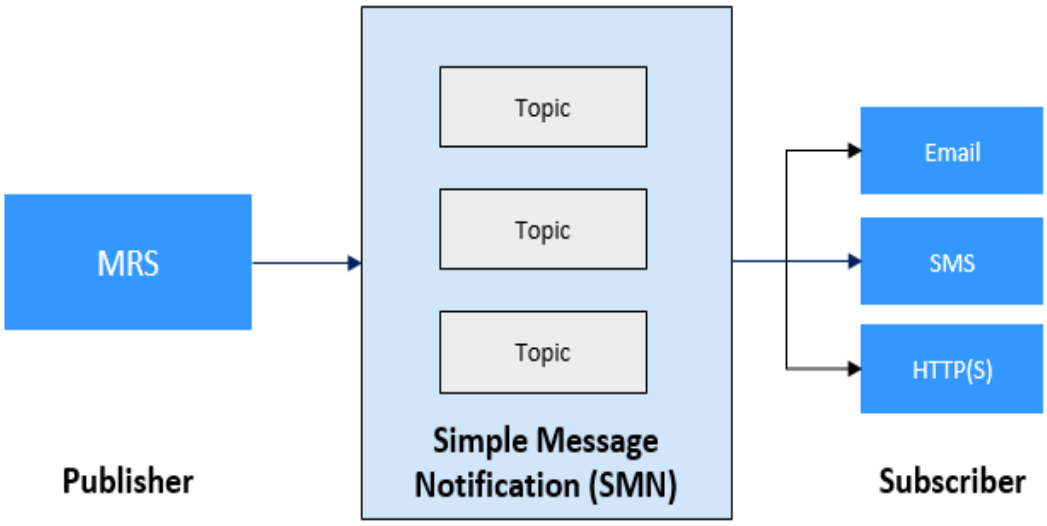
9.10.9 Configuring Notifications for MRS Cluster Alarms and Events

Simple Message Notification (SMN) is a reliable and flexible large-scale message notification service. It enables you to efficiently send messages to various

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

endpoints, such as email addresses and phone numbers. MRS uses SMN to offer a publish/subscribe model to achieve one-to-multiple message subscriptions and notifications through different subscription endpoints (SMSs and emails).

Figure 9-4 MRS cluster message notification



In an MRS cluster, you can choose whether to enable message notification in cluster alarm management to proactively notify users of specified messages.

- After a user subscribes to the notification service, the MRS management plane notifies the user of success or failure of manual cluster scale-out and scale-in, cluster deletion, and auto scaling by emails or SMS messages.
- The management plane checks the alarms about the MRS cluster and sends a notification to the tenant if the alarms are critical.
- If either of the operations such as deletion, shutdown, specifications modification, restart, and OS update is performed on an ECS in a cluster, the MRS cluster works abnormally. The management plane notifies a user when detecting that the VM of the user is in either of the preceding operations.

Step 1: Creating an SMN Topic and Adding a Subscription

A topic is used to publish messages and subscribe to notifications. It serves as a message sending channel, where publishers and subscribers can interact with each other.

Step 1 Log in to the management console.

Step 2 Click **Service List**. Under **Management & Governance**, click **Simple Message Notification**.

Step 3 In the navigation pane, choose **Topic Management** > **Topics**.

Step 4 Click **Create Topic**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-114 Parameters for creating an SMN topic

Parameter	Description	Example Value
Topic Name	Topic name, which: • Contains only letters, digits, hyphens (-), and underscores (_) and must start with a letter or digit. • Contains 1 to 255 characters. • Must be unique and cannot be modified after the topic is created.	mrs-job
Display name	Message sender name, which can contain up to 192 bytes. After you specify a display name, the sender in email messages will be presented as *Display name*<**username@example.com**>. Otherwise, the sender will be **username@example.com**.	mrs-job-failed
Enterprise Project	Name of an enterprise project. An enterprise project allows you to group cloud resources under specific projects, making it easier to manage cloud resources and users. The default name is **default**. This parameter is for enterprise users.	default

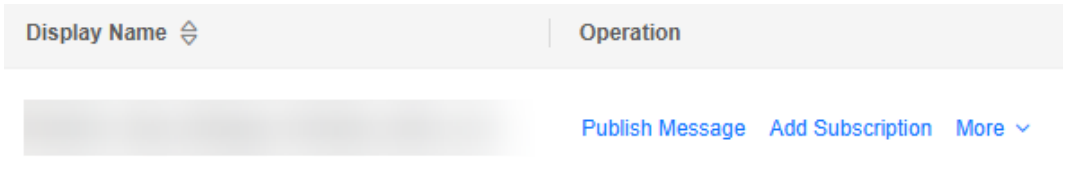
Step 5 Click **OK**.

After a topic is created, the system generates a topic URN, which uniquely identifies the topic and cannot be changed.

The created topic is displayed in the topic list. You can click the topic name to view its details.

Step 6 Locate the topic to which you want to add a subscription, click **More** in the **Operation** column, and select **Add Subscription**.

Figure 9-5 Adding a subscription to a topic



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Endpoint indicates the address of the subscription endpoint. You can add multiple SMS and email endpoints at a time, one in each line. You can enter a maximum of 10 endpoints.

Table 9-115 Parameters for adding a subscription

Parameter	Description	Example Value
Protocol	Protocol supported by the subscription endpoints. For example, you can receive messages through SMS messages or emails.	Email
Method	The following two options are available: • **Add subscriptions** • **Import subscriptions**: Select subscribers whose subscription has been confirmed and import the subscription data.	Add subscriptions
Endpoint	Subscription endpoint. You can enter up to 10 SMS, email, HTTP, or HTTPS endpoints, one in each line.	username@example.com

Step 7 Click **OK**. The new subscription is displayed in the subscription list.

Once you have added a subscription, SMN will send a confirmation message to the subscription endpoint, which includes a link for confirming the subscription. The subscription confirmation link is valid for 48 hours. Make sure to confirm your subscription on your mobile phone, mailbox, or other endpoints within this time frame.

The subscription endpoint can receive messages only after the subscription status changes to **Confirmed**.

----End

Configuring Notification Rules for MRS Clusters

Step 1 Log in to the MRS console.

Step 2 Choose **Active Clusters**, select a running cluster, and click its name to go to its details page.

Step 3 Click **Alarms**.

Step 4 Choose **Notification Rules** > **Add Notification Rule**.

Step 5 Set the notification rule parameters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-116 Parameters of a notification rule

Parameter	Example Value	Description
Rule Name	mrs_alarm	User-defined notification rule name. Only digits, letters, hyphens (-), and underscores (_) are allowed.
Message Notification	Enable	• If you enable this function, the system sends notifications to subscribers based on the notification rule. • If you disable this function, the rule does not take effect, that is, notifications are not sent to subscribers.
Topic Name	mrs_topic	Select an existing topic or click **Create Topic** to create a topic.
Notification Type	Alarm	Select the type of the notification to be subscribed to. • **Alarm** • **Event**
Subscription Items	Major	Select the items to be subscribed to. You can select all or some items as required. For example, choose **Major** > **Manager** and select **Node Fault**.

Step 6 Click **OK**. After the message notification is configured, the related endpoint can receive detailed notification messages when alarms or events in the subscription rule are generated in the MRS cluster.

----End

9.11 MRS Cluster Alarm Handling Reference

9.11.1 ALM-12001 Audit Log Dumping Failure

Description

Cluster audit logs need to be dumped on a third-party server due to the local historical data backup policy. The system starts to check the dump server at 3 a.m. every day. If the dump server meets the configuration conditions, audit logs can be successfully dumped. This alarm is generated when the audit log dump fails because the disk space of the dump directory on the third-party server is insufficient or a user changes the username, password, or dump directory of the dump server.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12001	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

System can store a maximum of only 50 dump files locally. If the fault persists on the dump server, the local audit logs may be lost, and the first 50 audit logs that exceed the current time cannot be queried.

Possible Causes

- The network connection is abnormal.
- The username, password, or dump directory of the dump server does not meet the configuration conditions.
- The disk space of the dump directory is insufficient.

Procedure

Check whether the network connection is normal.

Step 1 On the FusionInsight Manager home page, choose **Audit > Configurations**.

Step 2 Check whether the SFTP IP on the dump configuration page is valid.

Log in to the node where Manager is located as user **root** and run the **ping** command to check whether the network connection between the SFTP server and the cluster is normal.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 Repair the network connection, reset the SFTP password, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Wait for 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the username, password, or dump directory are correct.

Step 5 On the dump configuration page, check whether the username, password, and dump directory of the third-party server are correct.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Change the username, password, and dump directory, and click **OK** to deliver the configuration again.

Step 7 Wait for 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the disk space of the dump directory is sufficient.

Step 8 Log in to the third-party server as user **root** and run the **df** command to check whether the disk space of the dump directory of the third-party server exceeds 100 MB.

- If yes, go to **Step 11**.
- If no, go to **Step 9**.

Step 9 Expand disk space capacity for the third-party server, Reset the SFTP password and click **OK**

Step 10 Wait for 2 minutes, view real-time alarms and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Reset the dump rule.

Step 11 On the FusionInsight Manager home page, choose **Audit > Configurations**.

Step 12 Reset dump rules, set the parameters properly, and click **OK**.

Step 13 Wait for 2 minutes, view real-time alarms and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Collect fault information.

Step 14 On the FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 15 Select **OmmServer** from the **Service** and click **OK**.

Step 16 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 17 Contact the O&M personnel and send the collected log information.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.2 ALM-12004 OLdap Resource Abnormal

Description

The system checks LDAP resources every 60 seconds. This alarm is generated when the system detects that the LDAP resources in Manager are abnormal for six consecutive times.

This alarm is cleared when the Ldap resource in the Manager recovers and the alarm handling is complete.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12004	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Manager active/standby switchover may occur. The Manager and WebUI authentication service of components is unavailable. Security authentication and user management functions cannot be provided for upper-layer web services. As a result, you may fail to log in to the Manager and WebUI of components.

Possible Causes

The LdapServer process in the Manager is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check whether the LdapServer process in the Manager is normal.

Step 1 Log in the Manager node in the cluster as user **omm**.

Log in to FusionInsight Manager using the floating IP address, and run the **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** command to check the information about the current Manager two-node cluster.

Step 2 Run **ps -ef | grep slapd** command to check whether the LdapServer resource process in the **${BIGDATA_HOME}/om-server/om/** in the process configuration file is running properly.

NOTE

You can determine that the resource is normal by checking the following information:

1. After the **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** command runs, **ResHAStatus** of the OLdap is **Normal**.
2. After the **ps -ef | grep slapd** command runs, the slapd process of port 21750 can be viewed.
   - If yes, go to **Step 3**.
   - If no, go to **Step 4**.

Step 3 Run the **kill -2** *ldap pid* command and wait for 20 seconds. HA automatically starts the OLdap process. Check whether the current OLdap resource is in normal state.

- If yes, the operation is complete.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On the FusionInsight Manager home page, choose **O&M** > **Log** > **Download**.

Step 5 Select **OmsLdapServer** and **OmmServer** from the **Service** and click **OK**.

Step 6 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.3 ALM-12005 OKerberos Resource Abnormal

Description

The alarm module checks the status of the Kerberos resource in Manager every 80 seconds. This alarm is generated when the alarm module detects that the Kerberos resources are abnormal for six consecutive times.

This alarm is cleared when the Kerberos resource recovers and the alarm handling is complete.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12005	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The component WebUI authentication services are unavailable and cannot provide security authentication functions for web upper-layer services. Users may be unable to log in to FusionInsight Manager and the WebUIs of components.

Possible Causes

The OLdap resource on which the Okerberos depends is abnormal.

Procedure

Check whether the OLdap resource on which the Okerberos depends is abnormal in the Manager.

Step 1 Log in the Manager node in the cluster as user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Log in to FusionInsight Manager using the floating IP address, and run the **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** command to check the information about the current Manager two-node cluster.

Step 2 Run the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the OLdap resource managed by HA is normal. In single-node system, the OLdap resource is in the **Active_normal** state. In dual-node system, the OLdap resource is in the **Active_normal** state on the active node and in the **Standby_normal** state on the standby node.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 See the procedure in **ALM-12004 OLdap Resource Abnormal** to resolve the problem. After the OLdap resource status recovers, check whether the OKerberos resource status is normal.

- If yes, the operation is complete.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On the FusionInsight Manager home page, choose **O&M** > **Log** > **Download**.

Step 5 Select **OmsKerberos** and **OmmServer** from the **Service** and click **OK**.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.4 ALM-12006 Node Fault

Alarm Description

Controller checks the NodeAgent heartbeat every 30 seconds. If Controller does not receive heartbeat messages from a NodeAgent, it attempts to restart the NodeAgent process. This alarm is generated if the NodeAgent fails to be restarted for three consecutive times.

This alarm is cleared when Controller can properly receive the status report of the NodeAgent.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

In MRS 3.3.0 and later versions, the alarm name is changed to **NodeAgent Process Is Abnormal**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12006	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The NodeAgent process is abnormal and cannot report heartbeat messages to the platform. If the fault is caused by network or hardware faults or SSH mutual trust, component services may be abnormal.

Possible Causes

- The network is disconnected, the hardware is faulty, or the operating system runs slowly.
- The memory of the NodeAgent process is insufficient.
- The NodeAgent process is faulty.

Handling Procedure

Check whether the network is disconnected, whether the hardware is faulty, or whether the operating system runs commands slowly.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click ⌄ in the row containing the alarm, click the host name, and view address information about the host for which the alarm is generated.

Step 2 Log in to the active management node as user **root**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the faulty node is the active management node and fails login, the network of the active management node may be faulty. In this case, go to Step 4.

Step 3 Run the following command to check whether the faulty node is reachable:

```
ping IP address of the faulty host
```

- If yes, go to **Step 12**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Step 6 Contact the hardware administrator to check whether the hardware (CPU or memory) of the node is faulty.

- If yes, go to **Step 7**.
- If no, go to **Step 12**.

Step 7 Repair or replace faulty components and restart the node. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 8 If a large number of node faults are reported in the cluster, the floating IP addresses may be abnormal. As a result, Controller cannot detect the NodeAgent heartbeat.

Log in to any management node in the cluster as user **root** and view the **/var/log/Bigdata/omm/oms/ha/scriptlog/floatip.log** file to check whether the logs generated one to two minutes before and after the faults occur are complete.

```
cat /var/log/Bigdata/omm/oms/ha/scriptlog/floatip.log
```

For example, a complete log is in the following format:

```
2017-12-09 04:10:51,000 INFO (floatip) Read from ${BIGDATA_HOME}/om-server_*/om/etc/om/
routeSetConf.ini,value is : yes
2017-12-09 04:10:51,000 INFO (floatip) check wsNetExport : eth0 is up.
2017-12-09 04:10:51,000 INFO (floatip) check omNetExport : eth0 is up.
2017-12-09 04:10:51,000 INFO (floatip) check wsInterface : eRth0:oms, wsFloatIp: XXX.XXX.XXX.XXX.
2017-12-09 04:10:51,000 INFO (floatip) check omInterface : eth0:oms, omFloatIp: XXX.XXX.XXX.XXX.
2017-12-09 04:10:51,000 INFO (floatip) check  wsFloatIp : XXX.XXX.XXX.XXX is reachable.
2017-12-09 04:10:52,000 INFO (floatip) check  omFloatIp : XXX.XXX.XXX.XXX is reachable.
```

- If yes, go to **Step 12**.
- If no, go to **Step 9**.

Step 9 Check whether the omNetExport log is printed after the wsNetExport is detected or whether the interval for printing two logs exceeds 10 seconds or longer.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Check the **/var/log/message** file of the OS to determine whether sssd frequently restarts or nscd exception information is displayed when the fault occurs.

```
cat /var/log/message
```

Example of sssd restart information:

```
Feb  7 11:38:16 10-132-190-105 sssd[pam]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd[nss]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd[nss]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd[be[default]]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd: Starting up
Feb  7 11:38:16 10-132-190-105 sssd[be[default]]: Starting up
Feb  7 11:38:16 10-132-190-105 sssd[nss]: Starting up
Feb  7 11:38:16 10-132-190-105 sssd[pam]: Starting up
```

Example of the nscd exception information:

```
Feb 11 11:44:42 10-120-205-33 nscd: nss_ldap: failed to bind to LDAP server ldaps://10.120.205.55:21780: Can't contact LDAP server
Feb 11 11:44:43 10-120-205-33 ntpq: nss_ldap: failed to bind to LDAP server ldaps://10.120.205.55:21780: Can't contact LDAP server
Feb 11 11:44:44 10-120-205-33 ntpq: nss_ldap: failed to bind to LDAP server ldaps://10.120.205.92:21780: Can't contact LDAP server
```

- If yes, go to **Step 11**.
- If no, go to **Step 12**.

Step 11 Check whether the LdapServer node is faulty, for example, the service IP address is unreachable or the network latency is too long. If the fault occurs periodically, troubleshoot the fault when it occurs. Run the **top** command to check whether there is any software error.

To obtain information about the LdapServer node, log in to Manager (a cluster management tool), choose **Clusters** > **Services** > **LdapServer**, and click the **Instances** tab.

Check whether the memory of the NodeAgent process is insufficient.

Step 12 Log in to the faulty node as user **root** and run the following command to view the NodeAgent process logs:

```
vi /var/log/Bigdata/nodeagent/scriptlog/agent_gc.log.*.current
```

Step 13 Check whether the log file contains an error indicating that the metaspace size or heap memory size is insufficient.

Generally, the keyword for insufficient metaspace is **java.lang.OutOfMemoryError: Metaspace**, and the keyword for insufficient heap memory is **java.lang.OutOfMemoryError: Java heap space**.

The following is an example log containing an error:

```
java.lang.OutOfMemoryError:  Metaspace
        at java.lang.ClassLoader.defineClass1(Native Method)
        ...
```

- If yes, go to **Step 14**.
- If no, go to **Step 17**.

Step 14 Switch to user **omm**. Edit the corresponding file based on the cluster version, increase the values of **nodeagent.Xms** (NodeAgent initial heap memory) and **nodeagent.Xmx** (maximum heap memory), and save the modification.

```
su - omm
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The path of the file containing the parameters is as follows:

- For MRS clusters earlier than version 3.2.1:
  vi /opt/Bigdata/om-agent/nodeagent/bin/nodeagent_ctl.sh
- For MRS clusters of version 3.2.1 or later:
  vi $NODE_AGENT_HOME/etc/agent/nodeagent.properties

For example, change the value of **nodeagent.Xms** from **1024m** to **2048m**.

Step 15 Restart NodeAgent:

Command for stopping NodeAgent:

```
sh ${BIGDATA_HOME}/om-agent/nodeagent/bin/stop-agent.sh
```

Command for starting NodeAgent:

```
sh ${BIGDATA_HOME}/om-agent/nodeagent/bin/start-agent.sh
```

Step 16 Wait a moment and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 17**.

Check whether the NodeAgent process is faulty.

Step 17 Log in to the faulty node as user **root** and run the following command:

Command for switching to user **omm**:

```
su - omm
```

Command for viewing the NodeAgent process:

```
ps -ef | grep "Dprocess.name=nodeagent" | grep -v grep
```

Step 18 Check whether the NodeAgent process query command output is empty.

- If yes, go to **Step 19**.
- If no, go to **Step 21**.

Step 19 View the NodeAgent startup and run logs to locate the fault. After the fault is rectified, go to **Step 20**.

- NodeAgent run log:
  /var/log/Bigdata/nodeagent/agentlog/agent.log
- NodeAgent startup log:
  /var/log/Bigdata/nodeagent/scriptlog/nodeagent_ctl.log

Step 20 Restart the NodeAgent service:

Command for stopping NodeAgent:

```
sh ${BIGDATA_HOME}/om-agent/nodeagent/bin/stop-agent.sh
```

Command for starting NodeAgent:

```
sh ${BIGDATA_HOME}/om-agent/nodeagent/bin/start-agent.sh
```

Collect fault information.

Step 21 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 22 Select the following nodes from **Services** and click **OK**.

- NodeAgent
- Controller
- OS

Step 23 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 24 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.5 ALM-12007 Process Fault

Alarm Description

The process health check module checks the process status every 5 seconds. This alarm is generated when the process health check module detects that the process connection is faulty for three consecutive times.

This alarm is cleared when the process can be connected.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12007	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The impact varies depending on the instance that is faulty.

For example, if an HDFS instance is faulty, the impacts are as follows:

- If a DataNode instance is faulty, read and write operations cannot be performed on data blocks stored on the DataNode, which may cause data loss or unavailability. However, data in HDFS is redundant. Therefore, the client can access data from other DataNodes.
- If an HttpFS instance is faulty, the client cannot access files in HDFS over HTTP. However, the client can use other methods (such as shell commands) to access files in HDFS.
- If a JournalNode instance is faulty, namespaces and data logs cannot be stored to disks, which may cause data loss or unavailability. However, HDFS stores backups on other JournalNodes. Therefore, the faulty JournalNode can be recovered and data can be rebalanced.
- If a NameNode deployed in active/standby mode is faulty, an active/standby switchover occurs. If only one NameNode is deployed, the client cannot read or write any HDFS data. On MRS, NameNodes must be deployed in two-node mode.
- If a Router instance is faulty, the client cannot access data on the router. However, the client can use other Routers or directly access data on the backend NameNode.
- If a ZKFC instance is faulty, the NameNode does not continuously and automatically fail over. As a result, data cannot be read from or write to HDFS by the client. In this case, you need to enable automatic failover on other available ZKFC instances to restore the HDFS cluster.

Possible Causes

- The instance process is abnormal.
- The disk space is insufficient.

NOTE

If a large number of process fault alarms exist in a time segment, files in the installation directory may be deleted mistakenly or permission on the directory may be modified.

Handling Procedure

Check whether the instance process is abnormal.

Step 1 On FusionInsight Manager, choose **O&M > Alarm > Alarms**, click ⌄ in the row where the alarm is located, and click the host name to view the service IP address and service name of the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On the **Alarms** page, check whether the **ALM-12006 Node Fault** is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Handle the alarm according to **ALM-12006 Node Fault**.

Step 4 Log in to the host for which the alarm is generated as user **root**. Check whether the installation directory user, user group, and permission of the alarm role are correct. The user, user group, and the permission must be **omm:ficommon 750**.

For example, the NameNode installation directory is ***${BIGDATA_HOME}*/FusionInsight_Current/*1_8_NameNode*/etc**.

- If yes, go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Run the following command to set the permission to **750** and **User:Group** to **omm:ficommon**:

chmod 750 *<folder_name>*

chown omm:ficommon *<folder_name>*

Step 6 Wait for 5 minutes. In the alarm list, check whether **ALM-12007 Process Fault** is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Step 7 Log in to the active OMS node as user **root** and run the following command to view the **configurations.xml** file. In the preceding command, "Service name" is the service name queried in **Step 1**.

vi $BIGDATA_HOME/components/current/*Service name***/configurations.xml**

Search for the keyword **healthMonitor.properties**, find the health check configuration item corresponding to the alarm reporting instance, and record the interface or script path specified by **monitor.info**, as shown in the following figure.

Check the logs recorded in the interface or script and rectify the fault.

```
<config category="healthMonitor.properties" format="propertyfileconfigurer">
    <property type="hidden" scope="setup" classification="System">
        <name>monitor.type</name>
        <value>SCRIPT</value>
    </property>
    <property type="hidden" scope="setup" classification="System">
        <name>monitor.preInitDelay</name>
        <value>120000</value>
        <!-- 2min -->
    </property>
    <property type="hidden" scope="setup" classification="System">
        <name>monitor.recheckTimes</name>
        <value>90</value>
        </property>
    <property type="hidden" scope="setup" classification="System">
        <name>monitor.checkIntervals</name>
        <value>10000</value>
    </property>
    <property type="hidden" scope="setup" classification="System">
        <name>monitor.info</name>
        <value>#{install_home}/sbin/dbservice_mon.sh</value>
    </property>
    <property type="hidden" scope="setup" classification="System">
        <name>monitor.metric.connector</name>
        <value>#{install_home}/sbin/dbservice_ha.sh</value>
    </property>
</config>
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Wait for 5 minutes. In the alarm list, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Check whether the required disk is mounted on the faulty node normally.

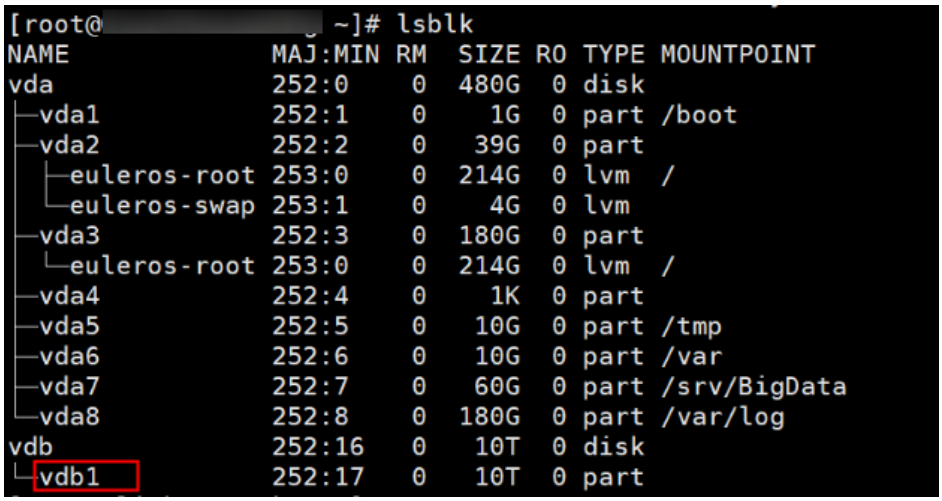
1. Log in to the node where the alarm was generated as the **root** user and run the following command to check whether there is a disk whose **TYPE** is **part** and **MOUNTPOINT** is empty.

   lsblk

   – If yes, the disk is not mounted. Go to **Step 9.2**.

   – If no, the disk is mounted. Go to **Step 11**.

   If information similar to the following is displayed, disk **vdb1** is not mounted.



2. Initialize the disk that fails to be mounted and mount it again.
3. Log in to the node for which the alarm was generated as the **root** user and run the following command to check and record the UUID of the mounted disk:

   blkid

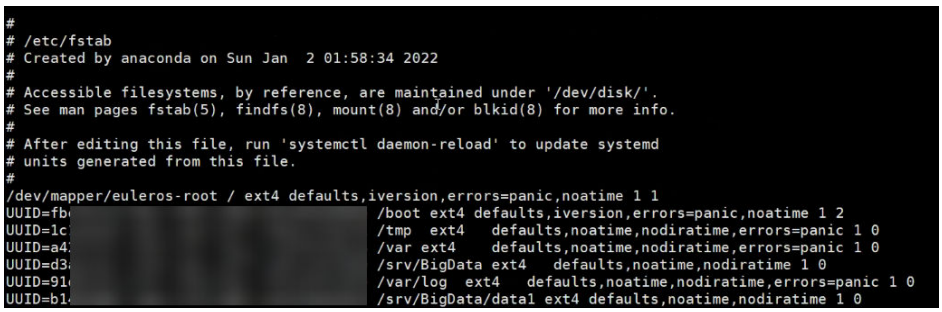
4. Run the following command to check whether there is the UUID of the mounted disk:

   vi /etc/fstab

   – If yes, go to **Step 9.6**.

   – If no, go to **Step 9.5**.

   The command output is shown in the following figure.



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

5. Add the mounted disk information to the **/etc/fstab** file. You can obtain the information in any of the following ways:
   - On the node where the alarm was generated, run the following command to view historical snapshot files and check the mounting information of the offline disk:

     cat /usr/local/diskmgt/ini/fstab.*xxx*
   - Obtain mounting information from the **/etc/fstab** file on a normal node in the current cluster.
   - Contact O&M personnel.
6. Restart the faulty instance on the node for which the alarm was generated.

   Log in to Manager, choose **Hosts**, and click the name of the host for which the alarm was generated. On the **Dashboard** tab page, at the **Instances** area, click a faulty instance. On the displayed instance page, choose **More** > **Restart Instance**. Repeat the preceding operations to restart all faulty instances on the host.

Step 10 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check whether disk space is sufficient.

Step 11 On the FusionInsight Manager, check whether the alarm list contains **ALM-12017 Insufficient Disk Capacity**.

- If yes, go to **Step 12**.
- If no, go to **Step 15**.

Step 12 Rectify the fault by following the steps provided in **ALM-12017 Insufficient Disk Capacity**.

Step 13 Wait for 5 minutes. In the alarm list, check whether **ALM-12017 Insufficient Disk Capacity** is cleared.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Wait for 5 minutes. In the alarm list, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Collect fault information.

Step 15 On the FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 16 According to the service name obtained in **Step 1**, select the component and **NodeAgent** from the **Service** and click **OK**.

Step 17 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 18 Contact the O&M personnel and send the collected log information.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.6 ALM-12010 Manager Heartbeat Interruption Between the Active and Standby Nodes

Alarm Description

This alarm is generated when the active Mager does not receive the heartbeat signal from the standby Manager within 7 seconds.

This alarm is cleared when the active Manager receives heartbeat signals from the standby Manager.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12010	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

When the active Manager process is abnormal, the active/standby switchover cannot be performed, affecting basic O&M functions.

Possible Causes

- The link between the active and standby Manager is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The node name configuration is incorrect.
- The port is disabled by the firewall.

Handling Procedure

Check whether the network between the active and standby Manager server is normal.

Step 1 In the FusionInsight Manager portal, click **O&M > Alarm > Alarms**, click ⌄ in the row containing the alarm and view the IP address of the standby Manager (Peer Manager) server in the alarm details.

Step 2 Log in to the active management node of the cluster as user **root**.

Step 3 Run the following command to check whether the standby FusionInsight Manager server is reachable:

```
ping heartbeat IP address of the standby FusionInsight Manager server
```

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the node name is correctly configured.

Step 6 Run the following command to go to the software installation directory of the active management node:

```
cd /opt
```

Step 7 Run the following command to find the configuration file directory of the active and standby nodes.

```
find -name hacom_local.xml
```

Step 8 Run the following command to go to the **workspace** directory:

```
cd ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/local/hacom/conf/
```

Step 9 Open the **hacom_local.xml** file. Check whether the local and peer nodes are correctly configured. The local node is configured as the active node, and the peer node is configured as the standby node.

```
vim hacom_local.xml
```

- If yes, go to **Step 12**.
- If no, go to **Step 10**.

Step 10 Modify the configuration of the active and standby nodes in the **hacom_local.xml** file and press **Esc** to return to the command mode. Run the **:wq** command to save the modification and exit.

Step 11 Check whether the alarm is cleared automatically.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check whether the port is disabled by the firewall.

Step 12 Run a command to check whether the heartbeat ports of the active and standby nodes are enabled. If the command output is displayed, the ports are enabled. Otherwise, the ports are disabled by the firewall.

```
lsof -i :20012
```

- If yes, go to **Step 13**.
- If no, go to **Step 16**.

Step 13 Run the following command to prevent disconnections from the server:

```
iptables -P INPUT ACCEPT
```

Step 14 Run the following command to clear the firewall:

```
iptables -F
```

Step 15 Check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 16**.

Collect fault information.

Step 16 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 17 Select the following nodes from the **Service** and click **OK**:

- OmmServer
- Controller
- NodeAgent

Step 18 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 19 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.7 ALM-12011 Manager Data Synchronization Exception Between the Active and Standby Nodes

Alarm Description

The system checks data synchronization between the active and standby Manager nodes every 60 seconds. This alarm is generated when the standby Manager fails to synchronize files with the active Manager.

This alarm is cleared when the standby Manager synchronizes files with the active Manager.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12011	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The configuration file of the standby Manager is not updated. When an active/standby switchover occurs, the configuration file that fails to be synchronized may be lost. Manager and some components may not run properly.

Possible Causes

- The link between the active and standby Managers is interrupted or The storage space of the **/srv/BigData/LocalBackup** directory is full.
- The synchronization file does not exist or the file permission is incorrect.

Handling Procedure

Check whether the network between the active Manager server and the standby Manager server is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 In the FusionInsight Manager portal, click **O&M > Alarm > Alarms**, click ⌄ in the row where the alarm is located and obtain the standby Manager server IP address (Peer Manager IP address) in the alarm details.

Step 2 Log in to the active OMS node of the cluster as user **root**.

You can view IP addresses of the active and standby OMS nodes on the MRS management console. The IP addresses are displayed on the **Nodes** tab page of the cluster details page.

Step 3 Check whether the standby Manager server is accessible:

```
ping IP address of the standby OMS node
```

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the storage space of the /srv/BigData/LocalBackup directory is full.

Step 6 Check whether the storage space of the **/srv/BigData/LocalBackup** directory is full:

```
df -hl /srv/BigData/LocalBackup
```

The command output is displayed as follows:

```
Filesystem      Size  Used Avail Use% Mounted on
/dev/vdb1        28G  1.6G   26G   6% /srv/BigData/LocalBackup
```

- If yes, go to **Step 7**.
- If no, go to **Step 10**.

Step 7 Clear unnecessary backup files:

```
rm -rf Directory to be cleared
```

The following is an example command:

```
rm -rf /srv/BigData/LocalBackup/0/default-oms_20191211143443
```

Step 8 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Backup Management**.

In the **Operation** column of the backup task to be performed, click **Configure** and change the value of **Maximum Number of Backup Copies** to reduce the number of backup file sets.

Step 9 Wait about 1 minute and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the synchronization file exists and whether the file permission is valid.

Step 10 Run the following command to check whether the synchronization file exists.

```
find /srv/BigData/ -name "sed*"
find /opt -name "sed*"
```

If no command output is displayed, there are no synchronization files.

- If yes, go to **Step 11**.
- If no, go to **Step 12**.

Step 11 Run the following command to view the synchronization file information and permission obtained in **Step 10**.

```
ll Path of the file you want to search for
```

- If the size of the file is 0 and the permission column is **-**, the file is a junk file. Run the following command to delete it.

  ```
  rm -rf File to be deleted
  ```

  Wait several minutes and check whether the alarm is cleared. If the alarm persists, go to **Step 12**.
- If the file size is not 0, go to **Step 12**.

Step 12 View the log files generated when the alarm is generated.

1. Run the following command to switch to the HA run log file path.

   ```
   cd /var/log/Bigdata/omm/oms/ha/runlog/
   ```

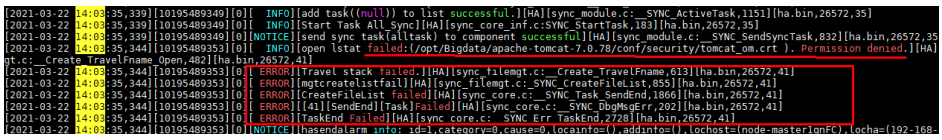
2. Decompress and view the log files generated when the alarm is generated.

   For example, if the name of the file to be viewed is **ha.log.2021-03-22_12-00-07.gz**, run the following command:

   ```
   gunzip ha.log.2021-03-22_12-00-07.gz
   vi ha.log.2021-03-22_12-00-07
   ```

   Check whether error information is reported before and after the alarm generation time.

   For example, if the following error information is displayed, the directory permission is insufficient. In this case, change the directory permission to be the same as that on the normal node.

   - If yes, rectify the fault based on the error information and go to **Step 13**.
   - If no, go to **Step 14**.

Step 13 Wait about 10 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Collect fault information.

Step 14 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 15 Select the following nodes from the **Service** and click **OK**:

- OmmServer

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Controller
- NodeAgent

Step 16 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 17 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.8 ALM-12012 NTP Service Is Abnormal

Alarm Description

The system checks whether the NTP service on a node synchronizes time with the NTP service on the active OMS node every 60 seconds. This alarm is generated when the NTP service fails to synchronize time for two consecutive times.

This alarm is generated when the time difference between the NTP service on a node and the NTP service on the active OMS node is greater than or equal to 20s for two consecutive times. This alarm is cleared when the time difference is less than 20s.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12012	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The time on the node is inconsistent with that on other nodes in the cluster. Therefore, some FusionInsight applications on the node may not run properly. If the time difference between the node and other Kerberos service instances keeps increasing, Kerberos authentication on the node may fail and service exceptions occur.

Possible Causes

- The NTP service on the current node cannot start properly.
- The current node fails to synchronize time with the NTP service on the active OMS node.
- The key authenticated by the NTP service on the current node is inconsistent with that on the active OMS node.
- The time offset between the node and the NTP service on the active OMS node is large.

Handling Procedure

Check the NTP service mode of the node.

Step 1 Log in to the active management node as the **root** user and check the resource status of the active and standby management nodes.

Switch to user **omm**:

```
su - omm
```

Check the resource status of the active and standby management nodes.

```
sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh
```

- If "chrony" is displayed in the **ResName** column of the command output, go to **Step 2**.
- If "ntp" is displayed in the **ResName** column, go to **Step 20**.

NOTE

If both "chrony" and "ntp" are displayed in the **ResName** column of the command output, the NTP service mode is being switched. Wait for 10 minutes and go to **Step 1** again. If both "chrony" and "ntp" persist, contact O&M personnel.

Check whether the chrony service on the node is started properly.

Step 2 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click ⌄ in the row containing the alarm, and view the name of the host for which the alarm is generated in **Location**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the chronyd process is running on the node where the alarm is generated. Log in to the node where the alarm is generated as user **root** and run the following command to check whether the chronyd process information is displayed:

```
ps -ef | grep chronyd | grep -v grep
```

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Start the NTP service.

Step 5 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the current node can synchronize time properly with the chrony service on the active OMS node.

Step 6 Check whether the node can synchronize time with the NTP service on the active OMS node based on additional information of the alarm.

- If yes, go to **Step 7**.
- If no, go to **Step 17**.

Step 7 Check whether the synchronization with the chrony service on the active OMS node is faulty.

Log in to the node for which the alarm is generated as user **root** and run the **chronyc sources** command.

```
chronyc sources
```

In the command output, if there is an asterisk (*) before the IP address of the chrony service on the active OMS node, the synchronization is normal. The command output is as follows:

```
MS Name/IP address         Stratum Poll Reach LastRx Last sample
===============================================================================
^* 10.10.10.162              10  10   377   626    +16us[  +15us] +/-  308us
```

In the command output, if there is no asterisk (*) before the IP address of the NTP service on the active OMS node, and the value of **Reach** is **0**, the synchronization is abnormal.

```
MS Name/IP address         Stratum Poll Reach LastRx Last sample
===============================================================================
^? 10.1.1.1                  0  10     0     -     +0ns[   +0ns] +/-    0ns
```

- If yes, go to **Step 8**.
- If no, go to **Step 38**.

Step 8 The chrony synchronization failure is typically caused by the system firewall. If the firewall can be disabled, disable it. If the firewall cannot be disabled, check the firewall configuration policy and ensure that UDP ports 123 and 323 are not disabled. (For details, see the firewall configuration policy of each system.)

Step 9 Check whether the alarm is cleared 10 minutes later.

- If yes, no further action is required.
- If no, go to **Step 10**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Log in to the active OMS node as user **root** and run the following command to view the authentication code whose key index is **1M**:

```
cat ${BIGDATA_HOME}/om-server/OMS/workspace/conf/chrony.keys
```

Step 11 Run the following command to check whether the key is the same as that queried in **Step 10**:

```
diff ${BIGDATA_HOME}/om-server/OMS/workspace/conf/chrony.keys /etc/chrony.keys
```

If the keys are the same, no result is returned after the command is executed. For example:

```
host01:~ # cat ${BIGDATA_HOME}/om-server/OMS/workspace/conf/chrony.keys
1 M sdYbq;o^CzEAWo<U=Tw5
host01:~ # diff ${BIGDATA_HOME}/om-server/OMS/workspace/conf/chrony.keys /etc/chrony.keys
host01:~ #
```

- If yes, go to **Step 12**.
- If no, go to **Step 38**.

Step 12 Run the following command to check whether the key is the same as that queried in **Step 10**: (Compare the key with that of the authentication key index **1M** queried in **Step 10**.)

```
cat ${BIGDATA_HOME}/om-server/om/packaged-distributables/ntpKeyFile
```

- If yes, go to **Step 13**.
- If no, go to **Step 15**.

Step 13 Log in to the faulty node as user **root** and run the **cat /etc/chrony.keys** command to check whether the key value is the same as that queried in **Step 12** (compare it with that of the authentication key index **1M**).

```
cat /etc/chrony.keys
```

- If yes, go to **Step 38**.
- If no, go to **Step 14**.

Step 14 Switch to user **omm**, change the key value of the authentication key index **1M** in **${NODE_AGENT_HOME}/chrony.keys** to the key value of **ntpKeyFile** in **Step 12**, and go to **Step 16**.

```
su - omm
vi ${NODE_AGENT_HOME}/chrony.keys
```

Step 15 Run the following commands as user **root** or **omm** to change the NTP key of the active OMS node (change **ntp.keys** to **ntpkeys** in Red Hat Enterprise Linux):

```
cd ${BIGDATA_HOME}/om-server/OMS/workspace/conf
sed -i "`cat chrony.keys | grep -n '1 M'|awk -F ':' '{print $1}'`d" chrony.keys
echo "1 M `cat ${BIGDATA_HOME}/om-server/om/packaged-distributables/ntpKeyFile`" >> chrony.keys
```

Check whether the key value of the authentication key index **1M** in **chrony.keys** is the same as that of **ntpKeyFile**.

- If yes, go to **Step 16**.
- If no, change the key of the authentication key index **1M** in **chrony.keys** to the key of **ntpKeyFile** and go to **Step 16**.

Step 16 After 5 minutes, restart the chrony service on the active OMS node. After 15 minutes, check whether the alarm is cleared.

```
systemctl restart chronyd
```

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 38**.

Check whether the time deviation between the node and the chrony service on the active OMS node is large.

Step 17 Check whether the time deviation is large in additional information of the alarm.

- If yes, go to **Step 18**.
- If no, go to **Step 38**.

Step 18 On the **Hosts** tab page, select the host for which the alarm is generated, and choose **More** > **Stop All Instances** to stop all the services on the node.

If the time on the alarm node is later than that on the chrony service of the active OMS node, adjust the time of the alarm node. After adjusting the time, choose **More** > **Start All Instances** to start the services on the node.

If the time on the alarm node is earlier than that on the chrony service of the active OMS node, wait until the time deviation is due and adjust the time of the alarm node. After adjusting the time, choose **More** > **Start All Instances** to start the services on the node.

NOTE

If you do not wait, data loss may occur.

Step 19 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 38**.

Check whether the NTP service on the node is started properly.

Step 20 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click ⌄ in the row containing the alarm, and view the name of the host for which the alarm is generated in **Location**.

Step 21 Check whether the ntpd process is running on the node using the following method. Log in to the node where the alarm is generated as user **root**. Run the following command to check whether the ntpd process information is displayed:

```
ps -ef | grep ntpd | grep -v grep
```

- If yes, go to **Step 24**.
- If no, go to **Step 22**.

Step 22 Start the NTP service:

```
service ntp start
```

For Red Hat operating systems, run the **service ntpd start** command.

Step 23 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 24**.

Check whether the node can synchronize time properly with the NTP service on the active OMS node.

Step 24 Check whether the node can synchronize time with the NTP service on the active OMS node based on additional information of the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 25**.
- If no, go to **Step 35**.

Step 25 Check whether the synchronization with the NTP service on the active OMS node is faulty.

Log in to the alarm node as user **root** and run the **ntpq -np** command.

```
ntpq -np
```

If an asterisk (*) exists before the IP address of the NTP service on the active OMS node in the command output, the synchronization is in normal state. The command output is as follows:

```
remote refid st t when poll reach delay offset jitter
==============================================================================
*10.10.10.162 .LOCL. 1 u 1 16 377 0.270 -1.562 0.014
```

If there is no asterisk (*) before the IP address of the NTP service on the active OMS node, as shown in the following command output, and the value of **refid** is **.INIT.**, the synchronization is abnormal.

```
remote refid st t when poll reach delay offset jitter
==============================================================================
10.10.10.162 .INIT. 1 u 1 16 377 0.270 -1.562 0.014
```

- If yes, go to **Step 26**.
- If no, go to **Step 38**.

Step 26 The NTP synchronization failure is typically caused by the system firewall. If the firewall can be disabled, run the **iptables -F** command to disable it. If the firewall cannot be disabled, run the **iptables -L** command to check the firewall configuration policy and ensure that the UDP port 123 is not disabled. (For details, see the firewall configuration policy of each system.)

```
iptables -F
```

Step 27 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 28**.

Step 28 Log in to the active OMS node as user **root** and run the following command to view the authentication key index **1M**:

```
cat ${BIGDATA_HOME}/om-server/OMS/workspace/conf/ntpkeys
```

Step 29 Run the following command to check whether the key is the same as that queried in **Step 28**:

```
diff ${BIGDATA_HOME}/om-server/OMS/workspace/conf/ntpkeys /etc/ntp/ntpkeys
```

If the keys are the same, no result is returned after the command is executed. For example:

```
host01:~ # cat ${BIGDATA_HOME}/om-server/OMS/workspace/conf/ntp.keys
1 M sdYbq;o^CzEAWo<U=Tw5
host01:~ # diff ${BIGDATA_HOME}/om-server/OMS/workspace/conf/ntp.keys /etc/ntp.keys
host01:~ #
```

- If yes, go to **Step 30**.
- If no, go to **Step 38**.

Step 30 Check whether the key value is the same as that queried in **Step 28**: (Compare the key with that of the authentication key index **1M** queried in **Step 28**.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
cat ${BIGDATA_HOME}/om-server/om/packaged-distributables/ntpKeyFile
```

- If yes, go to **Step 31**.
- If no, go to **Step 33**.

Step 31 Log in to the faulty node as user **root**. Check whether the key value is the same as that queried in **Step 30** (compare it with that of the authentication key index **1M**).

```
cat /etc/ntp/ntpkeys
```

- If yes, go to **Step 38**.
- If no, go to **Step 32**.

Step 32 Switch to user **omm**, change the key value of the authentication key index **1M** in **${NODE_AGENT_HOME}/ntp.keys** (**${NODE_AGENT_HOME}/ntpkeys** in Red Hat Enterprise Linux) to the key value of **ntpKeyFile** in **Step 30**, and go to **Step 34**.

```
su - omm
```

Step 33 Run the following commands as user **root** or **omm** to change the NTP key of the active OMS node (change **ntp.keys** to **ntpkeys** in Red Hat Enterprise Linux):

```
cd ${BIGDATA_HOME}/om-server/OMS/workspace/conf
sed -i "`cat ntp.keys | grep -n '1 M'|awk -F ':' '{print $1}'`d" ntp.keys
echo "1 M `cat ${BIGDATA_HOME}/om-server/om/packaged-distributables/ntpKeyFile`" >>ntp.keys
```

Check whether the key value of the authentication key index **1M** in **ntp.keys** is the same as that of **ntpKeyFile**.

- If yes, go to **Step 34**.
- If no, change the key of the authentication key index **1M** in **ntp.keys** to the key of **ntpKeyFile** and go to **Step 34**.

Step 34 After 5 minutes, restart the NTP service on the active OMS node. After 15 minutes, check whether the alarm is cleared.

```
service ntp restart
```

- If yes, no further action is required.
- If no, go to **Step 38**.

Check whether the time deviation between the node and the NTP service on the active OMS node is large.

Step 35 Check whether the time deviation is large in additional information of the alarm.

- If yes, go to **Step 36**.
- If no, go to **Step 38**.

Step 36 On the **Hosts** tab page, select the host for which the alarm is generated, and choose **More** > **Stop All Instances** to stop all the services on the node.

If the time on the alarm node is later than that on the NTP service of the active OMS node, adjust the time of the alarm node. After adjusting the time, choose **More** > **Start All Instances** to start the services on the node.

If the time on the alarm node is earlier than that on the NTP service of the active OMS node, wait until the time deviation is due and adjust the time of the alarm node. After adjusting the time, choose **More** > **Start All Instances** to start the services on the node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

If you do not wait, data loss may occur.

Step 37 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 38**.

Collect fault information.

Step 38 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 39 Expand the **Service** drop-down list, select **NodeAgent** and **OmmServer** for the target cluster, and click **OK**. Expand the **Hosts** dialog box and select the alarm node and the active OMS node.

Step 40 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 41 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.9 ALM-12014 Partition Lost

Alarm Description

The system checks the partition status every 60 seconds. This alarm is generated when the system detects that a partition to which service directories are mounted is lost (because the device is removed or goes offline, or the partition is deleted). The system checks the partition status periodically.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Clear
12014	Major	• Yes: MRS 3.3.0 and later versions • No: Versions earlier than MRS 3.3.0

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DirName	Specifies the directory for which the alarm is generated.
PartitionName	Specifies the device partition for which the alarm is generated.

Impact on the System

- Data loss: The device partition is lost and the data stored in the partition is lost.
- System breakdown: If the system disk is lost, the system deployed on the node cannot run properly. In some cases, the system may break down and cannot be started.
- Service failure: Read and write jobs on the lost device partition fail to run or run slowly.
- Service interruption: Customers may need time to restore data and systems, and services cannot be provided.
- Security risk: Important data may be stolen or disclosed, which severely affects customer services.

Possible Causes

- The hard disk is removed.
- The hard disk is offline, or a bad sector exists on the hard disk.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and click ⌄ in the row where the alarm is located.

Step 2 Obtain **HostName**, **PartitionName** and **DirName** from **Location**.

Step 3 Check whether the disk of **PartitionName** on **HostName** is inserted to the correct server slot.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Contact hardware engineers to remove the faulty disk.

Step 5 Log in to the **HostName** node where an alarm is reported and check whether there is a line containing **DirName** in the **/etc/fstab** file as user **root**.

```
cat /etc/fstab
```

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 Edit the file and delete the line containing **DirName**:

```
vi /etc/fstab
```

Step 7 Contact hardware engineers to insert a new disk. For details, see the hardware product document of the relevant model. If the faulty disk is in a RAID group, configure the RAID group. For details, see the configuration methods of the relevant RAID controller card.

Step 8 Wait 20 to 30 minutes (The disk size determines the waiting time), and run the **mount** command to check whether the disk has been mounted to the **DirName** directory.

- If yes, go to **Step 9** for MRS 3.3.0 and later, MRS 3.1.0.0.10/3.1.5.0.3 and later patch versions. For clusters earlier than MRS 3.3.0, manually clear the alarm. No further action is required.
- If no, go to **Step 10**.

Step 9 Wait about 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 11 Select the **OmmServer** from the Services drop-down list and click **OK**.

Step 12 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 13 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

MRS 3.3.0 and later, MRS 3.1.0.0.10/3.1.5.0.3 and later patch versions: After the fault is rectified, the system automatically clears this alarm.

Versions earlier than MRS 3.3.0: After the fault is rectified, the system does not automatically clear this alarm, and you need to manually clear the alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.10 ALM-12015 Partition Filesystem Readonly

Description

The system checks the partition status every 60 seconds. This alarm is generated when the system detects that a partition to which service directories are mounted enters the read-only mode (due to a bad sector or a faulty file system). The system checks the partition status periodically.

This alarm is cleared when the system detects that the partition to which service directories are mounted exits from the read-only mode (because the file system is restored to read/write mode, the device is removed, or the device is formatted).

Attribute

Alarm ID	Alarm Severity	Auto Clear
12015	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DirName	Specifies the directory for which the alarm is generated.
PartitionName	Specifies the device partition for which the alarm is generated.

Impact on the System

- Service failure: If a job needs to modify the data on the read-only device partition, the job may fail to run.
- Latency: If some components need to synchronize data to the read-only device partition, data synchronization may fail or time out, causing service delay.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The hard disk is faulty, for example, a bad sector exists.

Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and click ⌄ in the row that contains the alarm.

Step 2 Obtain **HostName** and **PartitionName** from **Location**. **HostName** is the node where the alarm is reported, and **PartitionName** is the partition of the faulty disk.

Step 3 Contact hardware engineers to check whether the disk is faulty. If the disk is faulty, remove it from the server.

Step 4 After the disk is removed, the system reports ALM-12014 Partition Lost. Handle the alarm by following the instructions in **ALM-12014 Partition Lost**. After the handling, the alarm is automatically cleared.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.11 ALM-12016 CPU Usage Exceeds the Threshold

Alarm Description

The system checks the CPU usage every 30 seconds and compares the actual CPU usage with the threshold. The CPU usage has a default threshold. This alarm is generated when the CPU usage exceeds the threshold for several times (configurable, 10 times by default) consecutively.

The alarm is cleared in the following two scenarios: The value of **Trigger Count** is 1 and the CPU usage is smaller than or equal to the threshold; the value of **Trigger Count** is greater than 1 and the CPU usage is smaller than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12016	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

- Latency: If the CPU usage of a host is too high, service processes may run slowly and services may be delayed.
- Service failure: If the host CPU usage is too high, service processing may slow down, time out, or fail. As a result, jobs may fail to run.

Possible Causes

- The alarm threshold or alarm smoothing times are incorrect.
- CPU configuration cannot meet service requirements. The CPU usage reaches the upper limit. Or the service is in peak hours. As a result, the CPU usage reaches the upper limit in a short period of time.

Handling Procedure

Check whether the alarm threshold or alarm Trigger Count are correct.

Step 1 Change the alarm threshold and alarm **Trigger Count** based on CPU usage.

On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Host** > **CPU** > **Host CPU Usage** and change the alarm smoothing times based on CPU usage.

Trigger Count indicates how many consecutive times the threshold is reached when the alarm is triggered.

On **Host CPU Usage** page and click **Modify** in the **Operation** column to change the alarm threshold.

Step 2 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check whether the CPU usage reaches the upper limit.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 In the alarm list on FusionInsight Manager, click ⌄ in the row where the alarm is located to view the alarm host address in the alarm details.

Step 4 On the **Hosts** page, click the node on which the alarm is reported.

Step 5 Observe the real-time data of the host CPU usage for about 5 minutes. If the CPU usage exceeds the threshold for multiple times, contact the MRS cluster administrator to increase the CPU specifications.

Step 6 Check whether the current traffic is in peak hours. If this alarm was generated during peak hours, expand the node capacity or contact the MRS cluster administrator to improve the CPU specifications.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On the FusionInsight Manager in the active cluster, choose **O&M** > **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, select **OmmServer**, and click **OK**.

Step 10 Click ✎ in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 11 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.12 ALM-12017 Insufficient Disk Capacity

Alarm Description

The system checks the host disk usage of the system every 30 seconds and compares the actual disk usage with the threshold. The disk usage has a default threshold, this alarm is generated when the host disk usage exceeds the specified threshold.

When the **Trigger Count** is 1, this alarm is cleared when the usage of a host disk partition is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the usage of a host disk partition is less than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12017	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the device partition for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If you need to modify or use data on the disk when the disk capacity is insufficient, the job may fail.

Possible Causes

- The alarm threshold is incorrect.
- The disk configuration cannot meet service requirements. The disk usage reaches the upper limit.

Handling Procedure

Check whether the threshold is set properly.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Host** > **Disk** > **Disk Usage** and check whether the threshold (configurable, 90% by default) is appropriate.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Locate the target threshold rule and click **Modify** in the **Operation** column to change the alarm threshold based on the current disk usage.

Step 3 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the disk usage reaches the upper limit.

Step 4 In the alarm list on FusionInsight Manager, click ⌄ in the row where the alarm is located to view the alarm host name and disk partition information in the alarm details.

Step 5 Log in to the node for which the alarm is generated as user **root**.

Step 6 Check the system disk partition usage. Check whether the disk is mounted to the following directories based on the disk partition name obtained in **Step 4**: **/**, **/opt**, **/tmp**, **/var**, **/var/log**, and **/srv/BigData** (can be customized).

```
df -lmPT |  awk '$2 != "iso9660"' | grep '^/dev/' | awk '{"readlink -m "$1 | getline real }{$1=real; print $0}' | sort -u -k 1,1
```

- If yes, the disk is a system disk. Then go to **Step 10**.
- If no, the disk is not a system disk. Then go to **Step 7**.

Step 7 Check the system disk partition usage. Determine the role of the disk based on the disk partition name obtained in **Step 4**.

```
df -lmPT |  awk '$2 != "iso9660"' | grep '^/dev/' | awk '{"readlink -m "$1 | getline real }{$1=real; print $0}' | sort -u -k 1,1
```

Step 8 Check whether the service that the disk belongs to is HDFS, Yarn, Kafka, Supervisor, or other services that require disk storage.

- If yes, expand the cluster capacity by referring to **Scaling Out an MRS Cluster** and go to **Step 9**.
- If no, go to **Step 12**.

Step 9 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Step 10 Check whether there is a file larger than 500 MB on the node: Check whether a large file that is written into the disk by mistake:

```
find / -xdev -size +500M -exec ls -l {} \;
```

- If yes, go to **Step 11**.
- If no, go to **Step 12**.

Step 11 Handle the large file and check whether the alarm is cleared 2 minutes later.

- If yes, no further action is required.
- If no, go to **Step 12**.

Step 12 Contact the system administrator to expand the disk capacity.

Step 13 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Collect fault information.

Step 14 On FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 15 Select **OMS** from the **Service** and click **OK**.

Step 16 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 17 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.13 ALM-12018 Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the system every 30 seconds and compares the actual memory usage with the threshold. The memory usage has a default threshold. This alarm is generated when the memory usage exceeds the threshold.

When the **Trigger Count** is 1, this alarm is cleared when the host memory usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the host memory usage is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12018	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

- Latency: If the host memory usage is too high, service processes may run slowly and services may be delayed.
- Service failure: If the host memory usage is too high, memory overflow may occur in service processes and jobs may fail.

Possible Causes

Memory configuration cannot meet service requirements. The memory usage reaches the upper limit.

Handling Procedure

Expand the system.

Step 1 Log in to FusionInsight Manager. In the alarm list, locate the alarm and click ⌄ to view the alarm host address.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 If the memory usage exceeds the preset threshold, increase the memory of the cluster.

Step 4 Check the system memory usage:

```
free -m | grep Mem\: | awk '{printf("%s,", $3 * 100 / $2)}'
```

Step 5 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager in the active cluster, choose **O&M** > **Log** > **Download**.

Step 7 Select **OmmServer** from the **Service** and click **OK**.

Step 8 Click ✎ in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.14 ALM-12027 Host PID Usage Exceeds the Threshold

Description

The system checks the PID usage every 30 seconds and compares the actual PID usage with the default PID usage threshold. This alarm is generated when the system detects that the PID usage exceeds the threshold.

When the **Trigger Count** is 1, this alarm is cleared when the PID usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the PID usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12027	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Service failure: If the PID usage of the host is too high, PIDs cannot be allocated to new service processes. As a result, jobs may fail to be executed.

Possible Causes

Too many processes are running on the node. You need to increase the value of **pid_max**.

Procedure

Increase the value of pid_max.

Step 1 In the alarm list on FusionInsight Manager, click ⌄ in the row where the alarm is located to view the alarm host address in the alarm details.

Step 2 Log in to the host where the alarm is generated as user **root**.

Step 3 Run the **cat /proc/sys/kernel/pid_max** command to check the value of **pid_max**.

Step 4 If the PID usage exceeds the threshold, edit the **/etc/sysctl.conf** file and increase the value of **kernel.pid_max** to twice of the value of **pid_max** queried in **Step 3**. If this parameter does not exist, add it to the end of the file.

For example, change the parameter to **kernel.pid_max=65536** and run the following command to make the parameter take effect immediately:

sysctl -p

NOTE

The maximum value of **kernel.pid_max** is as follows:

- On 32-bit systems: 32768
- On 64-bit systems: 4194304 (2^22)

Step 5 Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager home page of the active cluster, choose **O&M** > **Log** > **Download**.

Step 7 Select all services from the **Service** and click **OK**.

Step 8 Click ✎ in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.15 ALM-12028 Number of Processes in the D State and Z State on a Host Exceeds the Threshold

Alarm Description

The system checks the number of processes in the D state and Z state of user **omm** on the host every 30 seconds and compares the actual number with the threshold. The number of processes in the D state and Z state on the host has a default threshold range. This alarm is generated when the number of processes exceeds the threshold.

This alarm is cleared when **Trigger Count** is **1** and the total number of processes in the D state and Z state of user **omm** on the host does not exceed the threshold. This alarm is cleared when **Trigger Count** is greater than **1** and the total number of processes in the D state and Z state of user **omm** on the host is less than or equal to 90% of the threshold.

NOTE

The function of checking the number of processes in the Z state on the host applies to MRS 3.2.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12028	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

- Latency: New service processes cannot be created. Concurrent task processing may be slow and services may be delayed.
- Service failure: New service processes cannot be created, which may cause job failures.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state.

Handling Procedure

Check the processes in the D state and Z state.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the IP address of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Switch to user **omm**:

```
su - omm
```

Step 3 Run the following command as user **omm** to view the PID of the process that is in the D state and Z state:

```
ps -elf | grep -v "\[thread_checkio\]" | awk 'NR!=1 {print $2, $3, $4}' | grep omm | awk -F' ' '{print $1, $3}' | grep -E "Z|D" | awk '{print $2}'
```

Step 4 Check whether the command output is empty.

- If yes, the service process is running properly. Then go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Switch to user **root** and restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

```
reboot
```

Step 6 Check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Select **OMS** for **Service** and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.16 ALM-12033 Slow Disk Fault

Alarm Description

For MRS 3.3.0 and its later versions:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 1000 ms within 30 seconds in at least seven collection periods.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 150 ms within 300 seconds.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 1000 ms within 30 seconds in at least seven collection periods.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 20 ms within 300 seconds.

The collection period is 3 seconds, and the detection period is 30 or 300 seconds. This alarm is automatically cleared when neither of the preceding conditions is met for three consecutive detection periods (30 or 300 seconds).

For versions earlier than MRS 3.3.0:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - The system runs the **iostat** command every 3 seconds by default, and detects that the **svctm** value exceeds 1000 ms in at least seven consecutive periods within 30 seconds.
  - The system runs the **iostat** command every 3 seconds, and detects that more than 50% of I/Os take more than 150 ms within 300s.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - The system runs the **iostat** command every 3 seconds by default, and detects that the **svctm** value exceeds 1000 ms for at least 10 periods within 30 seconds.
  - The system runs the **iostat** command every 3 seconds by default, and detects that more than 60% of I/Os take more than 20 ms within 300s.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is automatically cleared when the preceding conditions have not been met for 15 minutes.

NOTE

For details about how to obtain the **svctm** value, see **Related Information**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12033	• Minor: MRS 3.3.0 and its later versions • Major: versions earlier than MRS 3.3.0	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
DiskName	Specifies the disk for which the alarm was generated.

Impact on the System

- The system I/O performance deteriorates, which means slow response and low throughput. For example, job submission is slow, page responds slowly, interface response times out, and the system is in error or even crash.
- System fault: Customer services may be interrupted. The system may break down and the key information stored on the faulty disk may be lost.

Possible Causes

The disk is aged or has bad sectors.

Handling Procedure

Check the disk status.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**.

Step 2 View the detailed information about the alarm. Check the values of **HostName** and **DiskName** in the location information to obtain the information about the faulty disk for which the alarm is generated.

Step 3 Check whether the node for which the alarm is generated is in a virtualization environment.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 Check whether the storage performance provided by the virtualization environment meets the hardware requirements. Then, go to **Step 5**.

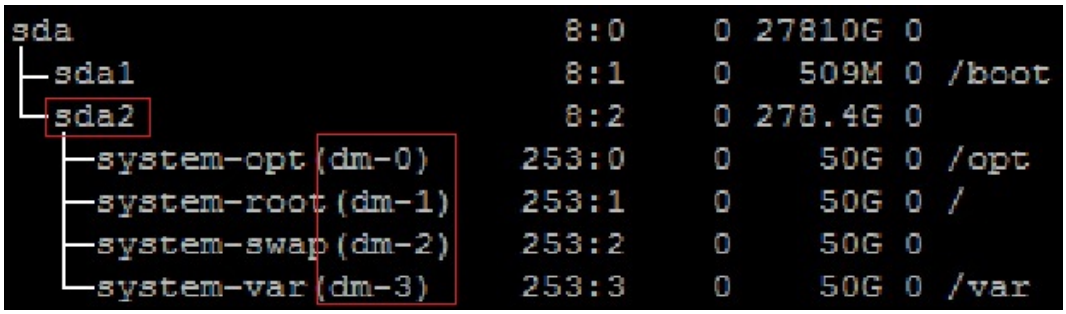
Step 5 Log in to the node where the alarm is generated as user **root**. Check whether the command output contains the value of the **DiskName** field.

```
df -h
```

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

Step 6 Check whether the mapping between the value of **DiskName** and the disk has been created.

```
lsblk
```



- If yes, go to **Step 7**.
- If no, go to **Step 22**.

Step 7 Log in to the node where the alarm is generated as user **root**. View the disk information and check whether RAID has been set up.

```
lsscsi | grep "/dev/sd[x]"
```

NOTE

In the command, **/dev/sd[x]** indicates the disk name obtained in **Step 2**.

Example:

```
lsscsi | grep "/dev/sda"
```

In the command output, if **ATA**, **SATA**, or **SAS** is displayed in the third line, the disk has not been organized into a RAID group. If other information is displayed, RAID has been set up.

- If yes, go to **Step 12**.
- If no, go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Run the **smartctl -i /dev/sd[x]** command to check whether the hardware supports the SMART tool.

Example:

```
smartctl -i /dev/sda
```

In the command output, if "SMART support is: Enabled" is displayed, the hardware supports SMART. If "Device does not support SMART" or other information is displayed, the hardware does not support SMART.

- If yes, go to **Step 9**.
- If no, go to **Step 16**.

Step 9 Run the **smartctl -H --all /dev/sd[x]** command to check basic SMART information and determine whether the disk is working properly.

Example:

```
smartctl -H --all /dev/sda
```

Check the value of **SMART overall-health self-assessment test result** in the command output. If the value is **FAILED**, the disk is faulty and needs to be replaced. If the value is **PASSED**, check the value of **Reallocated_Sector_Ct** or **Elements in grown defect list**. If the value is greater than 100, the disk is faulty and needs to be replaced.

- If yes, go to **Step 10**.
- If no, go to **Step 18**.

Step 10 Run the **smartctl -l error -H /dev/sd[x]** command to check the Glist of the disk and determine whether the disk is normal.

Example:

```
smartctl -l error -H /dev/sda
```

Check the **Command/Feature_name** column in the command output. If **READ SECTOR(S)** or **WRITE SECTOR(S)** is displayed, the disk has bad sectors. If other errors occur, the disk circuit board is faulty. Both errors indicate that the disk is abnormal and needs to be replaced.

If "No Errors Logged" is displayed, no error log exists. You can trigger the disk SMART self-check.

- If yes, go to **Step 11**.
- If no, go to **Step 18**.

Step 11 Run the **smartctl -t long /dev/sd[x]** command to trigger the disk SMART self-check. After the command is executed, the time when the self-check is to be completed is displayed. After the self-check is completed, repeat **Step 9** and **Step 10** to check whether the disk is working properly.

Example:

```
smartctl -t long /dev/sda
```

- If yes, go to **Step 17**.
- If no, go to **Step 18**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -H --all /dev/sd[x]** command to check whether the hardware supports SMART.

NOTE

- In the command, **[sat|scsi]** indicates the disk type. Both types need to be used.
- **[DID]** indicates the slot information. Slots 0 to 15 need to be used.

For example, run the following commands in sequence:

```
smartctl -d sat+megaraid,0 -H --all /dev/sda
smartctl -d sat+megaraid,1 -H --all /dev/sda
smartctl -d sat+megaraid,2 -H --all /dev/sda
...
```

Try the command combinations of different disk types and slot information. If "SMART support is: Enabled" is displayed in the command output, the disk supports SMART. Record the parameters of the disk type and slot information when a command is successfully executed. If "SMART support is: Enabled" is not displayed in the command output, the disk does not support SMART.

- If yes, go to **Step 13**.
- If no, go to **Step 16**.

Step 13 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -H --all /dev/sd[x]** command recorded in **Step 12** to check basic SMART information and determine whether the disk is normal.

Example:

```
smartctl -d sat+megaraid,2 -H --all /dev/sda
```

Check the value of **SMART overall-health self-assessment test result** in the command output. If the value is **FAILED**, the disk is faulty and needs to be replaced. If the value is **PASSED**, check the value of **Reallocated_Sector_Ct** or **Elements in grown defect list**. If the value is greater than 100, the disk is faulty and needs to be replaced.

- If yes, go to **Step 14**.
- If no, go to **Step 18**.

Step 14 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -l error -H /dev/sd[x]** command to check the Glist of the disk and determine whether the hard disk is working properly.

Example:

```
smartctl -d sat+megaraid,2 -l error -H /dev/sda
```

Check the **Command/Feature_name** column in the command output. If **READ SECTOR(S)** or **WRITE SECTOR(S)** is displayed, the disk has bad sectors. If other errors occur, the disk circuit board is faulty. Both errors indicate that the disk is abnormal and needs to be replaced.

If "No Errors Logged" is displayed, no error log exists. You can trigger the disk SMART self-check.

- If yes, go to **Step 15**.
- If no, go to **Step 18**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 15 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -t long /dev/sd[x]** command to trigger the disk SMART self-check. After the command is executed, the time when the self-check is to be completed is displayed. After the self-check is completed, repeat **Step 13** and **Step 14** to check whether the disk is working properly.

Example:

```
smartctl -d sat+megaraid,2 -t long /dev/sda
```

- If yes, go to **Step 17**.
- If no, go to **Step 18**.

Step 16 If the configured RAID controller card does not support SMART, the disk does not support SMART. In this case, use the check tool provided by the corresponding RAID controller card vendor to rectify the fault. Then go to **Step 17**.

For example, LSI is a MegaCLI tool.

Step 17 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, click **Clear** in the **Operation** column of the alarm, and check whether the alarm is reported on the same disk again.

If the alarm is reported for three times, replace the disk.

- If yes, go to **Step 18**.
- If no, no further action is required.

Replace the disk.

Step 18 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**.

Step 19 View the detailed information about the alarm. Check the values of **HostName** and **DiskName** in the location information to obtain the information about the faulty disk for which the alarm is reported.

Step 20 Replace the disk.

Step 21 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 22**.

Collect the fault information.

Step 22 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 23 Select **OMS** for **Service** and click **OK**.

Step 24 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 25 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

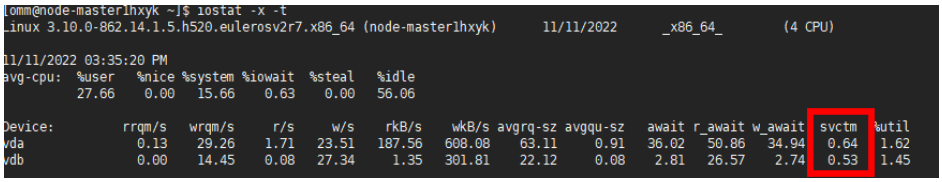
This alarm is automatically cleared after the fault is rectified.

Related Information

The **svctm** value can be obtained as follows:

- MRS 3.1.0:

  Run the **iostat -x -t** command in the OS.



- Versions later than MRS 3.1.0:

  svctm = (tot_ticks_new - tot_ticks_old) / (rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old)

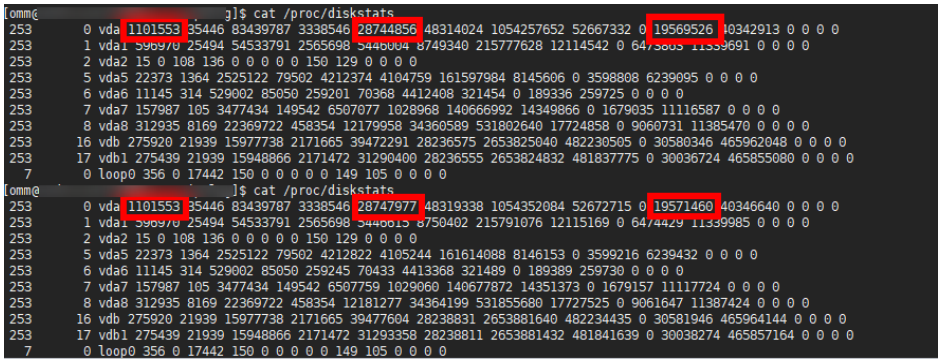
- Versions earlier than MRS 3.3.0: If **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, then **svctm = 0**.
- MRS 3.3.0 and its later versions:

  When the detection period is 30 seconds, if **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, then **svctm = 0**.

  When the detection period is 300 seconds and **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, if **tot_ticks_new - tot_ticks_old = 0**, then **svctm = 0**; otherwise, the value of **svctm** is infinite.

The parameters can be obtained as follows:

The system runs the **cat /proc/diskstats** command every 3 seconds to collect data. For example:



In the two commands:

In the data collected for the first time, the number in the fourth column is the **rd_ios_old** value, the number in the eighth column is the **wr_ios_old** value, and the number in the thirteenth column is the **tot_ticks_old** value.

In the data collected for the second time, the number in the fourth column is the **rd_ios_new** value, the number in the eighth column is the **wr_ios_new** value, and the number in the thirteenth column is the **tot_ticks_new** value.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In this case, the value of **svctm** is as follows:

(19571460 - 19569526) / (1101553 + 28747977 - 1101553 - 28744856) = 0.6197

9.11.17 ALM-12034 Periodical Backup Failure

Description

The system executes the periodic backup task every 60 minutes. This alarm is generated when a periodical backup task fails to be executed. This alarm is cleared when the next backup task is executed successfully.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12034	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
TaskName	Specifies the task.

Impact on the System

The periodic backup task failed, resulting in no available backup packages during the time period when the backup failed. When a system exception occurs and you need to use the backup package to restore data, no backup package is available during the failure period. As a result, data during the failure period cannot be restored.

Possible Causes

The alarm cause depends on the task details. Handle the alarm according to the logs and alarm details.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check whether the disk space is sufficient.

Step 1 In the FusionInsight Manager portal, click **O&M > Alarm > Alarms**.

Step 2 In the alarm list, click ⌄ in the row where the alarm is located and obtain **TaskName** from **Location**.

Step 3 Log in to the active node of the cluster as the **root** user and check the backup and restoration logs in **/var/log/Bigdata/controller/backup/**.

```
cd /var/log/Bigdata/controller/backup/
```

Run the **ll** command to view the log name. The following is an example of the command output:

```
[root@xxx backup]#ll
total 4476
-rw------- 1 omm wheel    4205 Jun 10 14:14 backupplugin_ctl.log
-rw------- 1 omm wheel  158640 Jun 11 00:02 backupplugin.log
-rw------- 1 omm wheel 4405518 Jun 11 10:28 controller_backup.log
```

Run the following command to check the backup and restoration logs:

```
vi Log file name
```

Check whether the log file contains information similar to the following:

```
Upload backup files to *** file failed, error info: ***
```

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On FusionInsight Manager, choose **O&M** > **Backup and Restoration > Backup Management**. Locate the backup task based on **TaskName**, click **Configure** in the **Operation** column, and check whether all configuration items are correctly configured.

- If yes, go to **Step 7**.
- If no, modify the configuration, save the modification, and go to **Step 5**.

Step 5 Choose **More > Back Up Now** to start the backup task and check whether the backup task is successfully executed.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Step 7 Choose **More > View History** in the **Operation** column. In the displayed dialog box view the task details.

Step 8 In the displayed dialog box and click ⌄ to check whether the following message is displayed: **Failed to backup xx due to insufficient disk space, move the data in the /srv/BigData/LocalBackup directory to other directories**.

- If yes, go to **Step 9**.
- If no, go to **Step 16**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Choose **Backup Path** > **View** and obtain the **Backup Path**.

Step 10 Log in to the node as user **root** and run the following command to view the node mounting details.

```
df -h
```

Step 11 Check whether the available space of the node to which the backup path is mounted is less than 20 GB.

- If yes, go to **Step 12**.
- If no, go to **Step 16**.

Step 12 Check whether there are many backup packages in the backup directory.

- If yes, go to **Step 13**.
- If no, go to **Step 16**.

Step 13 Remove the backup package from the backup directory or delete the backup package to ensure that the available space of the node to which the backup directory is mounted is greater than 20 GB.

Step 14 After the problem is resolved, perform the backup task again and check whether the backup task execution is successful.

- If yes, go to **Step 15**.
- If no, go to **Step 16**.

Step 15 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Collect fault information.

Step 16 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 17 Select **Controller** from the **Service** and click **OK**.

Step 18 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 19 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.18 ALM-12035 Unknown Data Status After Recovery Task Failure

Description

After the recovery task fails, the system automatically rolls back every 60 minutes. If the rollback fails, data may be lost. If this occurs, an alarm is reported. This alarm is cleared when the next recovery task execution is successful.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12035	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
TaskName	Specifies the task.

Impact on the System

After the recovery task fails, the system automatically rolls back. If the rollback fails, data may be lost or the data status may be unknown, which may affect services.

Possible Causes

The alarm cause depends on the task details. Handle the alarm according to the logs and alarm details.

Procedure

Collect fault information.

Step 1 In the FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services**, and check whether the running status of the component meets the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

requirements. (The OMS and DBService must be in the normal state, and other components must be stopped.)

- If yes, go to **Step 9**.
- If no, go to **Step 2**.

Step 2 Restore the component status as required and start the recovery task again.

Step 3 Log in to the FusionInsight Manager portal and click **O&M > Alarm > Alarms**.

Step 4 In the alarm list, click ⌄ in the row where the alarm is located to obtain **TaskName** from **Location**.

Step 5 Choose **O&M** > **Backup and Restoration > Restoration Management**.

Step 6 Find the restoration task by **Task Name** and view the task details.

Step 7 Perform the recovery task again and check whether the recovery task execution is successful.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 10 Select **Controller** from the **Service** and click **OK**.

Step 11 Click ✎ in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.19 ALM-12037 NTP Server Abnormal

Description

The system checks the NTP server status every 60 seconds. This alarm is generated when the system detects that the NTP server is abnormal for 10 consecutive times.

This alarm is cleared when the NTP server recovers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12037	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the IP address of the NTP server for which the alarm is generated.

Impact on the System

- The NTP server configured on the active OMS node is abnormal. In this case, the active OMS node cannot synchronize time with the NTP server and a time offset may be generated in the cluster.
- If the time difference exceeds 5 minutes, the client outside the cluster fails to authenticate the client inside the cluster. As a result, the job may fail to run.

Possible Causes

- The NTP server network is abnormal.
- The NTP server authentication fails.
- The NTP server time cannot be obtained.
- The time obtained from the NTP server is not continuously updated.

Procedure

Check the NTP server network.

Step 1 On the FusionInsight Manager portal, click **O&M > Alarm > Alarms** and click ⌄ in the row where the alarm is located.

Step 2 View the alarm additional information to check whether the NTP server fails to be pinged.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Contact the network administrator to check the network configuration and ensure that the network between the NTP server and the active OMS node is normal. Then, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the NTP server authentication fails.

Step 4 Log in to the active OMS node as user **root**.

Step 5 Run the following command to check the status of the resources on the active and standby nodes:

su - omm

sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh

- If "chrony" is displayed in the **ResName** column of the command output, go to **Step 6**.
- If "ntp" is displayed in the **ResName** column, go to **Step 7**.

NOTE

If both "chrony" and "ntp" are displayed in the **ResName** column of the command output, the NTP service mode is being switched. Wait for 10 minutes and perform **Step 5** again. If both "chrony" and "ntp" still exist in the **ResName** column, contact O&M personnel.

Step 6 Run the command **chronyc sources** to check whether the NTP server authentication fails.

If the value of **Reach** for chrony is **0**, the connection or authentication fails.

- If yes, go to **Step 12**.
- If no, go to **Step 8**.

Step 7 Run the command **ntpq -np** to check whether the NTP server authentication fails.

If **refid** of the NTP server is **.AUTH.**, the authentication fails.

- If yes, go to **Step 12**.
- If no, go to **Step 8**.

Check whether the time can be obtained from the NTP server.

Step 8 View the alarm additional information to check whether the time can be obtained from the NTP server.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Contact the provider of the NTP server to rectify the NTP server fault. After the NTP server is normal, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether the time obtained from the NTP server is not continuously updated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 View the alarm additional information to check whether the time obtained from the NTP server is not continuously updated.

- If yes, go to **Step 11**.
- If no, go to **Step 12**.

Step 11 Contact the provider of the NTP server to rectify the NTP server fault. After the NTP server is normal, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On the FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 13 Select **NodeAgent** and **OmmServer** from the **Service** and click **OK**.

Step 14 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 15 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.20 ALM-12038 Monitoring Indicator Dumping Failure

Description

After monitoring indicator dumping is configured on FusionInsight Manager, the system checks the monitoring indicator dumping result at the dumping interval (60 seconds by default). This alarm is generated when the dumping fails.

This alarm is cleared when dumping is successful.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12038	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The upper-layer management system cannot obtain monitoring indicators from the FusionInsight Manager system.

Possible Causes

- The server cannot be connected.
- The save path on the server cannot be accessed.
- The monitoring indicator file fails to be uploaded.

Procedure

Check whether the server connection is normal.

Step 1 Check whether the network between the FusionInsight Manager system and the server is normal.

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Contact the network administrator to recover the network and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Choose **System** > **Interconnection** > **Upload Performance Data** and check whether the FTP username, password, port, dump mode, and public key configured on the upload performance data page are consistent with the configuration on the server.

- If yes, go to **Step 5**.
- If no, go to **Step 4**.

Step 4 Enter the correct configuration information, click **OK**, and check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Check the permission of the save path on the server is correct.

Step 5 Choose **System** > **Interconnection > Upload Performance Data** and check the configuration items **FTP Username**, **Save Path**, and **Dump Mode**.

- If the dump mode is FTP, go to **Step 6**.
- If the dump mode is SFTP, go to **Step 7**.

Step 6 Log in to the server in FTP mode. In the default path, check whether **FTP Username** has the read and write permission of the relative path **Save Path**.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 7 Log in to the server in SFTP mode and check whether **FTP Username** has the read and write permission of the absolute path **Save Path**.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 8 Add the read and write permission and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check whether the save path on the server has sufficient disk space.

Step 9 Log in to the server and check whether the save path has sufficient disk space.

- If yes, go to **Step 11**.
- If no, go to **Step 10**.

Step 10 Delete unnecessary files or go to the monitoring indicator dumping configuration page to change the save path. Then, check whether the save path has sufficient disk space.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 12 Select **OMS** from the **Service** and click **OK**.

Step 13 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.21 ALM-12039 Active/Standby OMS Databases Not Synchronized

Description

The system checks the data synchronization status between the active and standby OMS Databases every 10 seconds. This alarm is generated when the synchronization status cannot be queried for 30 consecutive times or when the synchronization status is abnormal.

This alarm is cleared when the data synchronization status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12039	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Data on the active database cannot be synchronized to the standby database. If the active instance is abnormal during the alarm reporting period, service data may be lost or data on the FusionInsight Manager may be abnormal.

Possible Causes

- The network between the active and standby nodes is unstable.
- The standby OMS Database is abnormal.
- The standby node disk space is full.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check whether the network between the active and standby nodes is normal.

Step 1 Log in to FusionInsight Manager, click **O&M > Alarm > Alarms**, click ⌄ in the row where the alarm is located, and query the standby OMS Database IP address.

Step 2 Log in to the active OMS Database node as user **root**.

Step 3 Run the **ping** *Standby OMS Database heartbeat IP address* command to check whether the standby OMS Database node is reachable.

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the standby OMS database is normal.

Step 6 Log in to the standby OMS Database node as user **root**.

Step 7 Run the **su - omm** command to switch to user **omm**.

Step 8 Go to the **${BIGDATA_HOME}/om-server/om/sbin/** directory and run the **./status-oms.sh** command to check whether the standby OMS database resource status is normal. In the command output, check whether the following information is displayed in the row where **ResName** is **gaussDB**:

For example:

```
10_10_10_231 gaussDB Standby_normal Normal Active_standby
```

- If yes, go to **Step 9**.
- If no, go to **Step 16**.

Check whether the standby node disk space is full.(Skip this check for versions later than MRS 3.1.2.)

Step 9 Log in to the standby OMS Database node as user **root**.

Step 10 Run the **su - omm** command to switch to user **omm**.

Step 11 Run the **echo ${BIGDATA_DATA_HOME}/dbdata_om** command to obtain the OMS Database data directory.

Step 12 Run the **df -h** command to view the system disk partition usage information.

Step 13 Check whether the disk where the OMS Database data directory is mounted is full.

- If yes, go to **Step 14**.
- If no, go to **Step 16**.

Step 14 Expand the disk capacity.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 15 After the disk capacity is expanded, wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Collect fault information.

Step 16 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 17 Select **OMMServer** from the **Service** and click **OK**.

Step 18 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 19 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.22 ALM-12040 Insufficient System Entropy

Alarm Description

MRS 3.2.0 or later:

The system checks whether the rng-tools or haveged tool has been enabled and correctly configured every 5 minutes. If neither tool is configured, this alarm is generated. If either is configured, the system continues to check the entropy. If the entropy is less than 100 for five consecutive times, this alarm is generated.

This alarm is cleared when rng-tools or haveged has been installed and enabled on the target node and the entropy of the OS is greater than or equal to 100 in at least one of five entropy checks.

MRS 3.1.2 or earlier:

The system checks the entropy for five consecutive times at 00:00 every day. Specifically, the system checks whether rng-tools or haveged has been enabled and correctly configured. If neither is configured, the system continues to check the entropy. If the entropy is less than 100 for five consecutive times, this alarm is reported.

This alarm is cleared when the system detects that the true random number mode has been configured, the random number parameters have been configured in the pseudo-random number mode, or neither mode is configured but the entropy of the OS is greater than or equal to 100 in at least one of five entropy checks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12040	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The execution of encryption and decryption commands on the node may be slow. As a result, the service processing performance of each instance deteriorates, and even service processes cannot run.

Possible Causes

- rng-tools or haveged has not been installed or started.
- The entropy of the OS is smaller than 100 for multiple consecutive times.

Handling Procedure

Check whether haveged or rng-tools has been installed or started.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**.

Step 2 Check the value of **HostName** in the **Location** area to obtain the name of the host for which the alarm is generated.

Step 3 Log in to the node for which the alarm is generated as user **root**.

Step 4 Run the following command to check the haveged installation status and check whether the command output is empty:

```
/bin/rpm -qa | grep -w "haveged"
```

- If yes, go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Check the command output:

```
/sbin/service haveged status |grep "running"
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the command is executed successfully, haveged has been installed and configured correctly and is running properly. Go to **Step 8**.
- If the command fails to execute, haveged is not running properly. Run the following command to manually restart haveged and go to **Step 9**:
  ```
  systemctl restart haveged.service
  ```

Step 6 Run the following command to check the rng-tools installation status and check whether the command output is empty:

```
/bin/rpm -qa | grep -w "rng-tools"
```

- If yes, contact the OS vendor to install and start haveged or rng-tools. Then go to **Step 9**.
- If no, go to **Step 7**.

Step 7 Check the command output:

```
ps -ef | grep -v "grep" | grep rngd  | tr -d " " | grep "\-r/dev/urandom"
```

- If the command is executed successfully, rngd has been installed and configured correctly and is running properly. Go to **Step 8**.
- If the command fails to execute, rngd is not running properly. Run the following command to manually restart rngd and go to **Step 9**:
  ```
  systemctl restart rngd.service
  ```

Check the entropy of the OS.

Step 8 Manually check the entropy of the OS.

Log in to the target node as user **root** and check whether the entropy of the OS meets cluster installation requirements (no less than 100).

```
cat /proc/sys/kernel/random/entropy_avail
```

- If yes, the entropy of the OS is not less than 100. Go to **Step 9**.
- If no, the entropy of the OS is less than 100. Use either of the following methods and go to **Step 9**.
  - Method 1: Use haveged (true random number mode). Contact the OS vendor to install and start haveged.
  - Method 2: Use rng-tools (pseudo-random number mode). Contact the OS vendor to install and start rng-tools and configure it based on the OS type.

Step 9 Wait till the system checks the entropy again. Check whether the alarm is cleared automatically.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Select **NodeAgent** for **Service** and click **OK**.

Step 12 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

9.11.23 ALM-12041 Incorrect Permission on Key Files

Alarm Description

The system checks whether the permission, user, and user group information about critical directories or files is normal every 5 minutes. This alarm is generated when the information is abnormal.

This alarm is cleared when the information becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12041	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
PathName	Specifies the path or name of the abnormal file.

Impact on the System

System functions are unavailable.

- If the permission of key files **okerberos** and **oldap** is abnormal, the authentication fails and the job may fail to run.
- If the permission of key files **controller** and **pms** is abnormal, the process becomes faulty, which may affect the elastic scaling performance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the permission on key Tomcat files is abnormal, the login and viewing functions of FusionInsight Manager are affected.

Possible Causes

The file permission is abnormal or the file is lost due to a user manually modified information such as the file permission, user, and user group, or the system is powered off unexpectedly.

Handling Procedure

Check whether the abnormal file exists and whether the permission on the abnormal file is correct.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**.

Step 2 Check the value of **HostName** to obtain the host name involved in this alarm. Check the value of **PathName** to obtain the path or name of the abnormal file.

Step 3 Log in to the node for which the alarm is generated as user **root**.

Step 4 Run the following command to check the current user, permissions, and user group information of the file or directory on the host:

```
ll Path name
```

The path name is the abnormal file path obtained in **Step 2**.

Step 5 Go to **${BIGDATA_HOME}/om-agent/nodeagent/etc/agent/autocheck** directory. Then search for the name of the abnormal file and check the due permissions on the file.

```
cd ${BIGDATA_HOME}/om-agent/nodeagent/etc/agent/autocheck
vi keyfile
```

NOTE

To ensure proper configuration synchronization between the active and standby OMS servers, files, directories, and files and sub-directories in the directories configured in **$OMS_RUN_PATH/workspace/ha/module/hasync/plugin/conf/filesync.xml** will also be monitored except files and directories in **keyfile**. User **omm** must have read and write permissions of files and read and execute permissions of directories.

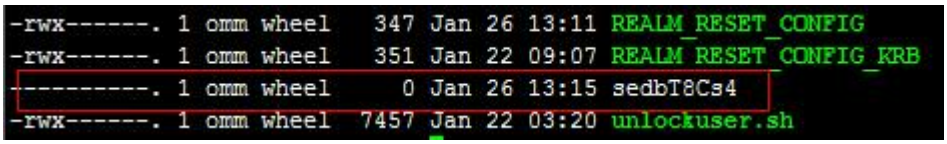
Step 6 Compare the real-world permission of the file with the due permission obtained in **Step 5** and correct the permission, user, and user group information for the file.

Step 7 Wait an hour and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

If the disk partition where the cluster installation directory resides is used up, some temporary files will be generated in the program installation directory when running the **sed** command fails. Users do not have the read, write, and execute permissions of these temporary files. The system reports an alarm indicating that permissions of temporary files are abnormal if these files are within the monitoring range of the alarm. Perform the preceding alarm handling processes to clear the alarm. Alternatively, you can directly delete the temporary files after confirming that the files with abnormal permissions are temporary. The following is an example temporary file generated after a **sed** command execution failure.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



Collect fault information.

Step 8 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 9 Expand the **Service** drop-down list, select **NodeAgent**, and click **OK**.

Step 10 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 11 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.24 ALM-12042 Incorrect Configuration of Key Files

Description

The system checks whether critical configurations are correct every 5 minutes. This alarm is generated when the configurations are abnormal.

This alarm is cleared when the configurations become normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12042	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
PathName	Specifies the path or name of the abnormal file.

Impact on the System

Functions related to the file are abnormal.

- If the configuration of key files **okerberos** and **oldap** is abnormal, the authentication fails and the job may fail to run.
- If the permission on the controller and pms key files is abnormal, the process may be faulty, which may affect the elastic scaling performance.
- If the permission on key Tomcat files is abnormal, the login and viewing functions of FusionInsight Manager are affected.

Possible Causes

The file configuration is modified manually or the system is powered off unexpectedly.

Procedure

Check abnormal file configuration.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**.

Step 2 Check the value of **HostName** to obtain the host name involved in this alarm. Check the value of **PathName** to obtain the path or name of the abnormal file.

Step 3 Log in to the node for which the alarm is generated as user **root**.

Step 4 View the $BIGDATA_LOG_HOME/nodeagent/scriptlog/checkfileconfig.log file and analyze the cause based on the error log. Locate the check standards of the file in the **Related Information** and manually check and modify the file based on the standards.

Run the **vi** *file name* command to enter the editing mode, and then press **Insert** to start editing.

After the modification is complete, press **Esc** to exit the editing mode and enter **:wq** to save the settings and exit.

For example:

vi /etc/ssh/sshd_config

Step 5 Wait an hour and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **NodeAgent** from the **Service** and click **OK**.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

- **Check standards of /etc/fstab**

  Check whether the partitions configured in the **/etc/fstab** file can be found in **/proc/mounts**.

  Check whether the swap partitions configured in fstab correspond to those in /proc/swaps.

- **Check the /etc/hosts configuration file.**

  Run **cat /ect/hosts**. If any of the following situations occurs, the **/etc/hosts** configuration file is abnormal:

  a. The **/etc/hosts** file does not exist.
  b. The host name is not configured in the file.
  c. The host name maps to multiple IP addresses in the file.
  d. The IP address corresponding to the host name does not exist in the command output of the **ifconfig** command.
  e. One IP address maps to multiple host names in the file.
  f. The file does not contain the Hadoop local domain name mapping, for example, xxx hadoop.example.com. (applicable only to MRS 3.2.0-LTS.1.10)

- **Check standards of /etc/ssh/sshd_config**

  Run the **vi /etc/ssh/sshd_config** command to check whether configuration items are configured as follows:

  a. The value of **UseDNS** must be set to **no**.
  b. The value of **MaxStartups** must be greater than or equal to 1000.
  c. At least one of the **PasswordAuthentication** and **ChallengeResponseAuthentication** parameters must be left blank or at least one of the parameters must be set to **yes**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.25 ALM-12045 Read Packet Dropped Rate Exceeds the Threshold

Alarm Description

The system checks the read packet dropped rate every 30 seconds. This alarm is generated when the read packet dropped rate exceeds the threshold (the default threshold is 0.5%) for multiple times (the default value is **5**).

To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Packet Dropped Rate**.

This alarm is cleared when **Trigger Count** is 1 and the read packet dropped rate is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the read packet dropped rate is less than or equal to 90% of the threshold.

The alarm detection is disabled by default. If you want to enable this function, check whether this function can be enabled based on Checking System Environments.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12045	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
PortName	Specifies the network port for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- Latency: Requests are responded slowly and services are delayed.
- Service failure: Requests cannot be responded or time out. Jobs may fail to run.

Risk warning: In SUSE kernel 3.0 or later or Red Hat 7.2, the system kernel modifies the mechanism for counting the number of dropped read packets. In this case, this alarm may be generated even if the network is running properly, but services are not affected. You are advised to check the system environment first.

Possible Causes

- The NICs are bonded in active/standby mode.
- The alarm threshold is improperly configured.
- The network quality is poor.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click ⌄ in the row containing the alarm, and view the name of the host for which the alarm is generated and the NIC name.

Check whether the NICs are bonded in active/standby mode.

Step 2 Log in to the alarm node as user **omm** and run the **ls -l /proc/net/bonding** command to check whether the **/proc/net/bonding** directory exists on the node.

- If yes, the bond mode is configured for the node. Go to **Step 3**.

```
# ls -l /proc/net/bonding/
total 0
-r--r--r-- 1 root root 0 Oct 11 17:35 bond0
```

- If no, the bond mode is not configured for the node. Go to **Step 5**.

```
# ls -l /proc/net/bonding/
ls: cannot access /proc/net/bonding/: No such file or directory
```

Step 3 Run the **cat /proc/net/bonding/***bond0* command to check whether the value of **Bonding Mode** in the configuration file is **fault-tolerance**.

NOTE

In the command, **bond0** indicates the name of the bond configuration file. Use the file name obtained in **Step 2**.

```
# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v3.7.1 (April 27, 2011)

Bonding Mode: fault-tolerance (active-backup)
Primary Slave: eth1 (primary_reselect always)
Currently Active Slave: eth1
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 0
Down Delay (ms): 0

Slave Interface: eth0
MII Status: up
Speed: 1000 Mbps
Duplex: full
Link Failure Count: 1
Slave queue ID: 0
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
Slave Interface: eth1
MII Status: up
Speed: 1000 Mbps
Duplex: full
Link Failure Count: 1
Slave queue ID: 0
```

- If yes, the NICs are configured in active/standby bond mode. Go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Check whether the NIC specified by **NetworkCardName** in the alarm is the standby NIC.

- If yes, the alarm of the standby NIC cannot be automatically cleared. Manually clear the alarm on the alarm management page. No further action is required.
- If no, go to **Step 5**.

NOTE

To determine the standby NIC, check the **/proc/net/bonding/bond0** configuration file. If the NIC name corresponding to **NetworkCardName** is **Slave Interface** but not **Currently Active Slave** (the current active NIC), the NIC is the standby one.

Check whether the threshold is set properly.

Step 5 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Packet Dropped Rate**, and check whether the alarm threshold is configured properly. The default value is **0.5%**. You can adjust the threshold as needed.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Packet Dropped Rate**. Click **Modify** in the **Operation** column to change the threshold.

Step 7 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the network connection is normal.

Step 8 Contact the network administrator to check whether the network is normal.

- If yes, rectify the fault and go to **Step 9**.
- If no, go to **Step 10**.

Step 9 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect the fault information.

Step 10 On FusionInsight Manager of the active cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Select **OMS** for **Service** and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Expand the **Hosts** dialog box and select the alarm node and the active OMS node.

Step 13 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.26 ALM-12046 Write Packet Dropped Rate Exceeds the Threshold

Alarm Description

The system checks the write packet dropped rate every 30 seconds. This alarm is generated when the write packet dropped rate exceeds the threshold (the default threshold is 0.5%) for multiple times (the default value is **5**).

To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Packet Dropped Rate**.

If **Trigger Count** is **1**, this alarm is cleared when the network write packet dropped rate is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the network write packet dropped rate is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12046	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Port Name	Specifies the network port for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

- Latency: Requests are responded slowly and services are delayed.
- Service failure: Requests cannot be responded or time out. Jobs may fail to run.

Possible Causes

- The alarm threshold is improperly configured.
- The network quality is poor.

Handling Procedure

Check whether the threshold is set properly.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Packet Dropped Rate**, and check whether the alarm threshold is configured properly. The default value is **0.5%**. You can adjust the threshold as needed.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Packet Dropped Rate**. Click **Modify** in the **Operation** column to change the threshold.

Step 3 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network connection is normal.

Step 4 Contact the network administrator to check whether the network is normal.

- If yes, rectify the fault and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 After 5 minutes, check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect the fault information.

Step 6 On FusionInsight Manager of the active cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Select **OMS** for **Service** and click **OK**.

Step 8 Expand the **Hosts** dialog box and select the alarm node and the active OMS node.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.27 ALM-12047 Read Packet Error Rate Exceeds the Threshold

Alarm Description

The system checks the read packet error rate every 30 seconds. This alarm is generated when the read packet error rate exceeds the threshold (the default threshold is **0.5%**) for multiple times (the default value is **5**).

To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Packet Error Rate**.

If **Trigger Count** is **1**, this alarm is cleared when the read packet error rate is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the read packet error rate is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12047	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Port Name	Specifies the network port for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

- Latency: Requests are responded slowly and services are delayed.
- Service failure: Requests cannot be responded or time out. Jobs may fail to run.

Possible Causes

- The alarm threshold is improperly configured.
- The network quality is poor.

Handling Procedure

Check whether the threshold is set properly.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Packet Error Rate**, and check whether the alarm threshold is configured properly. The default value is **0.5%**. You can adjust the threshold as needed.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Packet Error Rate**. Click **Modify** in the **Operation** column to change the threshold.

Step 3 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network connection is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Contact the network administrator to check whether the network is normal.

- If yes, rectify the fault and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect the fault information.

Step 6 On FusionInsight Manager of the active cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Select **OMS** for **Service** and click **OK**.

Step 8 Expand the **Hosts** dialog box and select the alarm node and the active OMS node.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.28 ALM-12048 Write Packet Error Rate Exceeds the Threshold

Alarm Description

The system checks the write packet error rate every 30 seconds. This alarm is generated when the write packet error rate exceeds the threshold (the default threshold is **0.5%**) for multiple times (the default value is **5**).

To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Packet Error Rate**.

If **Trigger Count** is **1**, this alarm is cleared when the write packet error rate is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the write packet error rate is less than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12048	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Port Name	Specifies the network port for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

- Latency: Requests are responded slowly and services are delayed.
- Service failure: Requests cannot be responded or time out. Jobs may fail to run.

Possible Causes

- The alarm threshold is improperly configured.
- The network quality is poor.

Handling Procedure

Check whether the threshold is set properly.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Packet Error Rate**, and check whether the alarm threshold is configured properly. The default value is **0.5%**. You can adjust the threshold as needed.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Packet Error Rate**. Click **Modify** in the **Operation** column to change the threshold.

Step 3 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network connection is normal.

Step 4 Contact the network administrator to check whether the network is normal.

- If yes, rectify the fault and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect the fault information.

Step 6 On FusionInsight Manager of the active cluster, choose **O&M** > **Log** > **Download**.

Step 7 Select **OMS** for **Service** and click **OK**.

Step 8 Expand the **Hosts** dialog box and select the alarm node and the active OMS node.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.29 ALM-12049 Network Read Throughput Rate Exceeds the Threshold

Description

The system checks the network read throughput rate every 30 seconds and compares the actual throughput rate with the threshold (the default threshold is 80%). This alarm is generated when the system detects that the network read throughput rate exceeds the threshold for several times (5 times by default) consecutively.

To change the threshold, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Throughput Rate**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

When the **Trigger Count** is 1, this alarm is cleared when the network read throughput rate is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the network read throughput rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12049	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

- Latency: When the host network read throughput exceeds the threshold, the request response slows down, causing service delay.
- Service failure: When the host network read throughput exceeds the threshold, requests cannot be properly responded or timed out, which may cause job execution failures.

Possible Causes

- The alarm threshold is set improperly.
- The network port rate cannot meet the current service requirements.

Procedure

Check whether the threshold is set properly.

Step 1 On the FusionInsight Manager, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Throughput Rate** and

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

check whether the alarm threshold is set properly. (By default, 80% is a proper value. However, users can configure the value as required.)

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Based on actual usage condition, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Reading** > **Read Throughput Rate** and click **Modify** in the **Operation** column to modify the alarm threshold.

Step 3 Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network port rate can meet the service requirements.

Step 4 On FusionInsight Manager, click ⌄ in the row where the alarm is located in the real-time alarm list and obtain the IP address of the host and the network port name for which the alarm is generated.

Step 5 Log in to the host for which the alarm is generated as user **root**.

Step 6 Run the **ethtool** *network port name* command to check the maximum speed of the current network port.

NOTE

In the VM environment, you cannot run a command to query the network port rate. It is recommended that you contact the system administrator to confirm whether the network port rate meets the requirements.

Step 7 If the network read throughput rate exceeds the threshold, contact the system administrator to increase the network port rate.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager home page of the active cluster, choose **O&M** > **Log > Download**.

Step 10 Select **OMS** from the **Service** and click **OK**.

Step 11 Set **Host** to the node for which the alarm is generated and the active OMS node.

Step 12 Click ✎ in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.30 ALM-12050 Network Write Throughput Rate Exceeds the Threshold

Description

The system checks the network write throughput rate every 30 seconds and compares the actual throughput rate with the threshold (the default threshold is 80%). This alarm is generated when the system detects that the network write throughput rate exceeds the threshold for several times (5 times by default) consecutively.

To change the threshold, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Throughput Rate**.

When the **Trigger Count** is 1, this alarm is cleared when the network write throughput rate is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the network write throughput rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12050	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- Latency: When the host network write throughput exceeds the threshold, the request response is slowed down and services are delayed.
- Service failure: When the host network write throughput exceeds the threshold, requests cannot be responded or times out, which may cause job running failures.

Possible Causes

- The alarm threshold is set improperly.
- The network port rate cannot meet the current service requirements.

Procedure

Check whether the threshold is set properly.

Step 1 On the FusionInsight Manager, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Throughput Rate** and check whether the alarm threshold is set properly. (By default, 80% is a proper value. However, users can configure the value as required.)

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Based on actual usage condition, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Writing** > **Write Throughput Rate** and click **Modify** in the **Operation** column to modify the alarm threshold.

Step 3 Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network port rate can meet the service requirements.

Step 4 On FusionInsight Manager, click ⌄ in the row where the alarm is located in the real-time alarm list and obtain the IP address of the host and the network port name for which the alarm is generated.

Step 5 Log in to the host for which the alarm is generated as user **root**.

Step 6 Run the **ethtool** *network port name* command to check the maximum speed of the current network port.

NOTE

In the VM environment, you cannot run a command to query the network port rate. It is recommended that you contact the system administrator to confirm whether the network port rate meets the requirements.

Step 7 If the network write throughput rate exceeds the threshold, contact the system administrator to increase the network port rate.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Collect fault information.

Step 9 On the FusionInsight Manager home page of the active cluster, choose **O&M** > **Log > Download**.

Step 10 Select **OMS** from the **Service** and click **OK**.

Step 11 Set **Host** to the node for which the alarm is generated and the active OMS node.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.31 ALM-12051 Disk Inode Usage Exceeds the Threshold

Alarm Description

The system checks the disk Inode usage every 30 seconds and compares the actual Inode usage with the threshold (the default threshold is 80%). This alarm is generated when the Inode usage exceeds the threshold for several times (5 times by default) consecutively.

To change the threshold, choose **O&M > Alarm** > **Thresholds** > **Host** > **Disk** > **Disk Inode Usage**.

When the **Trigger Count** is 1, this alarm is cleared when the disk Inode usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the disk Inode usage is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12051	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the disk partition for which the alarm is generated.
Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

Service failure: If you need to modify or use data on the disk when data cannot be written to the file system, the job may fail.

Possible Causes

Massive small files are stored in the disk.

Handling Procedure

Massive small files are stored in the disk.

Step 1 On FusionInsight Manager, choose **O&M > Alarm > Alarms** and click ⌄ in the row where the alarm is located in the real-time alarm list and obtain the IP address of the host and the disk partition for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the following command to check the disk Inode usage:

```
df -i | grep -iE Partition name|Filesystem
```

The command output is as follows:

```
# df -i | grep -iE "xvda2|Filesystem"
Filesystem            Inodes   IUsed   IFree IUse% Mounted on
/dev/xvda2            2359296  207420 2151876    9% /
```

Step 4 If the Inode usage exceeds the threshold, manually check small files stored in the disk partition and confirm whether these small files can be deleted.

Run the **for i in /*; do echo $i; find $i|wc -l; done** command to query the number of files in a partition. Replace **/*** with the specified partition.

```
# for i in /srv/*; do echo $i; find $i|wc -l; done
/srv/BigData
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
4284
/srv/ftp
1
/srv/www
13
```

- If yes, delete the files and go to **Step 5**:
  rm -rf *Path of the files or folder to be deleted*

  NOTICE

  Deleting a file or folder is a high-risk operation. Ensure that the file or folder is no longer required before performing this operation.

- If no, expand the disk capacity. Then, perform **Step 5**.

Step 5 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager homepage of the active cluster, choose **O&M** > **Log** > **Download**.

Step 7 Select **OMS** from the **Service** and click **OK**.

Step 8 Set **Host** to the node for which the alarm is generated and the active OMS node.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.32 ALM-12052 TCP Temporary Port Usage Exceeds the Threshold

Description

The system checks the TCP temporary port usage every 30 seconds and compares the actual usage with the threshold (the default threshold is 80%). This alarm is generated when the TCP temporary port usage exceeds the threshold for several times (5 times by default) consecutively.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

To change the threshold, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Network Status** > **TCP Ephemeral Port Usage**.

When the **Trigger Count** is 1, this alarm is cleared when the TCP temporary port usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the TCP temporary port usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12052	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

Services on the host may fail to establish connections, causing service interruption.

Possible Causes

- The temporary port cannot meet the current service requirements.
- The system is abnormal.

Procedure

Expand the temporary port number range.

Step 1 On FusionInsight Manager, click ⌄ in the row where the alarm is located in the real-time alarm list and obtain the IP address of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Run the **cat /proc/sys/net/ipv4/ip_local_port_range |cut -f 1** command to obtain the value of the start port and run the **cat /proc/sys/net/ipv4/ip_local_port_range |cut -f 2** command to obtain the value of the end port. The total number of temporary ports is the value of the end port minus the value of the start port. If the total number of temporary ports is smaller than 28,232, the random port range of the OS is narrow. Contact the system administrator to increase the port range.

Step 4 Run the **ss -ant 2>/dev/null | grep -v LISTEN | awk 'NR > 2 {print $4}'| awk -F':' '{print $NF}' | awk '$1 >"***Value of the start port***" {print $1}' | sort -u | wc -l** command to calculate the number of used temporary ports.

Step 5 The formula for calculating the usage of the temporary ports is: Usage of the temporary ports = (Number of used temporary ports/Total number of temporary ports) x 100%. Check whether the temporary port usage exceeds the threshold.

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

Step 6 Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check whether the system environment is abnormal.

Step 7 Run the following command to import the temporary file and view the frequently used ports in the **port_result.txt file**:

netstat -tnp|sort > $BIGDATA_HOME/tmp/port_result.txt

```
netstat -tnp|sort

Active Internet connections (w/o servers)

Proto Recv Send LocalAddress ForeignAddress State PID/ProgramName tcp   0   0 10-120-85-154:45433
10-120-85-154:9866 CLOSE_WAIT 94237/java
tcp   0   0 10-120-85-154:45434  10-120-85-154:9866 CLOSE_WAIT 94237/java
tcp   0   0 10-120-85-154:45435  10-120-85-154:9866 CLOSE_WAIT 94237/java
...
```

Step 8 Run the following command to view the processes that occupy a large number of ports:

ps -ef |grep *PID*

NOTE

- PID is the processes ID queried in **Step 7**.
- Run the following command to collect information about all processes and check the processes that occupy a large number of ports:

  ps -ef > $BIGDATA_HOME/tmp/ps_result.txt

Step 9 After obtaining the administrator's approval, clear the processes that occupy a large number of ports. Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 On the FusionInsight Manager home page of the active cluster, choose **O&M** > **Log** > **Download**.

Step 11 Select **OMS** from the **Service** and click **OK**.

Step 12 Set **Host** to the node for which the alarm is generated and the active OMS node.

Step 13 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected log information and files **port_result.txt** and **ps_result.txt**. Then, delete the two residual temporary files from the environment.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.33 ALM-12053 Host File Handle Usage Exceeds the Threshold

Description

The system checks the file handle usage every 30 seconds and compares the actual usage with the threshold (the default threshold is 80%). This alarm is generated when the host file handle usage exceeds the threshold for several times (5 times by default) consecutively.

To change the threshold, choose **O&M > Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Host Status** > **Host File Handle Usage**.

When the **Trigger Count** is 1, this alarm is cleared when the host file handle usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the host file handle usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12053	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

Service failure: When the host file handle usage exceeds the threshold, system applications cannot perform I/O operations such as file opening and network operations. As a result, the program is abnormal, which may cause job running failure.

Possible Causes

- The application process is abnormal. For example, the opened file or socket is not closed.
- The number of file handles cannot meet the current service requirements.
- The system is abnormal.

Procedure

Check information about files opened in processes.

Step 1 On FusionInsight Manager, click ⌄ in the row where the alarm is located in the real-time alarm list and obtain the IP address of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the following command to check the process that occupies excessive file handles.

for proc in /proc/[0-9]*; do if [-d "$proc/fd"]; then num_fds=$(ls -l "$proc/fd" | wc -l); pid=$(basename $proc); echo "$num_fds ${pid}" ; fi; done | sort -nr | more

Step 4 Check whether the processes in which a large number of files are opened are normal. For example, check whether there are files or sockets not closed.

- If yes, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 7**.

Step 5 Release the abnormal processes that occupy too many file handles.

Step 6 Five minutes later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Increase the number of file handles.

Step 7 On FusionInsight Manager, click ⌄ in the row where the alarm is located in the real-time alarm list and obtain the IP address of the host for which the alarm is generated.

Step 8 Log in to the host for which the alarm is generated as user **root**.

Step 9 Contact the system administrator to increase the number of system file handles.

Step 10 Run the **cat /proc/sys/fs/file-nr** command to view the used handles and the maximum number of file handles. The first value is the number of used handles, the third value is the maximum number. Please check whether the usage exceeds the threshold.

```
# cat /proc/sys/fs/file-nr
12704    0    640000
```

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 11 Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check whether the system environment is abnormal.

Step 12 Contact the system administrator to check whether the operating system is abnormal.

- If yes, rectify the operating system fault and go to **Step 13**.
- If no, go to **Step 14**.

Step 13 Wait for 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Collect fault information.

Step 14 On the FusionInsight Manager home page of the active cluster, choose **O&M** > **Log** > **Download**.

Step 15 Select **OMS** from the **Service** and click **OK**.

Step 16 Set **Host** to the node for which the alarm is generated and the active OMS node.

Step 17 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 18 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.34 ALM-12054 Invalid Certificate File

Alarm Description

The system checks whether the certificate file is invalid (has expired or is not valid yet) on 23:00 every day. This alarm is generated when the certificate file is invalid.

This alarm is cleared when a valid certificate is imported and the alarm detection mechanism is triggered on the next hour.

NOTE

For MRS 3.2.0 or later, the certificate file is checked at the beginning of each hour.

For versions earlier than MRS 3.2.0, the certificate file is checked on 23:00 every day.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12054	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The functions of the related modules cannot be used.

Possible Causes

No certificate (CA certificate, HA root certificate, HA user certificate, Gaussdb root certificate, or Gaussdb user certificate) is imported to the system, the certificate fails to be imported, or the certificate file is invalid.

Handling Procedure

Check the alarm cause.

Step 1 On FusionInsight Manager, locate the target alarm in the real-time alarm list and click ⌄.

View **Additional Information** to obtain the additional information about the alarm.

- If **CA Certificate** is displayed in the additional alarm information, log in to the active OMS management node as user **omm** and go to **Step 2**.
- If **HA root Certificate** is displayed in the additional information, view **Location** to obtain the name of the host involved in this alarm. Then, log in to the host as user **omm** and go to **Step 3**.
- If **HA server Certificate** is displayed in the additional information, view **Location** to obtain the name of the host involved in this alarm. Then, log in to the host as user **omm** and go to **Step 4**.
- If **Certificate has expired** is displayed in the additional information, view **Location** to obtain the name of the host for which the alarm is generated. Then, log in to the host as user **omm** and perform **Step 2** to **Step 4** in sequence to check whether the certificates have expired. If these certificates have not expired, check whether other certificates have been imported. If yes, import the certificate files again.

Check the validity period of the certificate files in the system.

Step 2 Check whether the current system time is in the validity period of the CA certificate.

Run the **bash ${CONTROLLER_HOME}/security/cert/conf/querycertvalidity.sh** command to check the effective time and due time of the CA root certificate.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

Step 3 Check whether the current system time is in the validity period of the HA root certificate.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/root-ca.crt** command to check the effective time and due time of the HA root certificate.

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Check whether the current system time is in the validity period of the HA user certificate.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/server.crt** command to check the effective time and due time of the HA user certificate.

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

The following is an example of the effective time and due time of a CA or HA certificate:

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            97:d5:0e:84:af:ec:34:d8
        Signature Algorithm: sha256WithRSAEncryption
        Issuer: C=CN, ST=xxx, L=yyy, O=zzz, OU=IT, CN=HADOOP.COM
        Validity
            Not Before: Dec 13 06:38:26 2016 GMT          // Effective time
            Not After : Dec 11 06:38:26 2026 GMT          // Due time
```

Import certificate files.

Step 5 Import a new CA certificate file.

Apply for or generate a new CA certificate file and import it. The alarm is automatically cleared after the CA certificate is imported. Check whether this alarm is reported again during periodic check.

- If yes, go to **Step 7**.
- If no, no further action is required.

Step 6 Import a new HA certificate file.

Apply for or generate a new HA certificate file and import it. The alarm is automatically cleared after the CA certificate is imported. Check whether this alarm is reported again during periodic check.

- If yes, go to **Step 7**.
- If no, no further action is required.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 In the **Services** area, select **Controller**, **OmmServer**, **OmmCore**, and **Tomcat**, and click **OK**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.35 ALM-12055 Certificate File Is About to Expire

Alarm Description

The system checks the certificate file on 23:00 every day. This alarm is generated if the certificate file is about to expire within 30 days.

This alarm is cleared when a certificate that is not about to expire is imported and the alarm detection mechanism is triggered on the next hour.

NOTE

For MRS 3.2.0 or later, the certificate file is checked at the beginning of each hour.

For versions earlier than MRS 3.2.0, the certificate file is checked on 23:00 every day.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12055	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Some functions will be unavailable if the certificate expires.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The remaining validity period of a system certificate (CA certificate, HA root certificate, HA user certificate, Gaussdb root certificate, or Gaussdb user certificate) is less than 30 days.

Handling Procedure

Check the alarm cause.

Step 1 On FusionInsight Manager, locate the target alarm in the real-time alarm list and click ⌄.

View **Additional Information** to obtain the additional information about the alarm.

- If **CA Certificate** is displayed in the additional alarm information, log in to the active OMS management node as user **omm** and go to **Step 2**.
- If **HA root Certificate** is displayed in the additional information, view **Location** to obtain the name of the host involved in this alarm. Then, log in to the host as user **omm** and go to **Step 3**.
- If **HA server Certificate** is displayed in the additional information, view **Location** to obtain the name of the host involved in this alarm. Then, log in to the host as user **omm** and go to **Step 4**.

Check the validity period of the certificate files in the system.

Step 2 Check whether the remaining validity period of the CA certificate is smaller than the alarm threshold.

Run the **bash ${CONTROLLER_HOME}/security/cert/conf/querycertvalidity.sh** command to check the effective time and due time of the CA root certificate.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 3 Check whether the remaining validity period of the HA root certificate is smaller than the alarm threshold.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/root-ca.crt** command to check the effective time and due time of the HA root certificate.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 4 Check whether the remaining validity period of the HA user certificate is smaller than the alarm threshold.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/server.crt** command to check the effective time and due time of the HA user certificate.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The following is an example of the effective time and due time of a CA or HA certificate:

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            97:d5:0e:84:af:ec:34:d8
        Signature Algorithm: sha256WithRSAEncryption
        Issuer: C=CN, ST=xxx, L=yyy, O=zzz, OU=IT, CN=HADOOP.COM
        Validity
            Not Before: Dec 13 06:38:26 2016 GMT            // Effective time
            Not After : Dec 11 06:38:26 2026 GMT            // Due time
```

Import certificate files.

Step 5 Import a new CA certificate file.

Apply for or generate a new CA certificate file and import it. Manually clear the alarm and check whether this alarm is generated again during periodic check.

- If yes, go to **Step 7**.
- If no, no further action is required.

Step 6 Import a new HA certificate file.

Apply for or generate a new HA certificate file and import it. Manually clear the alarm and check whether this alarm is generated again during periodic check.

- If yes, go to **Step 7**.
- If no, no further action is required.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 In the **Services** area, select **Controller**, **OmmServer**, **OmmCore**, and **Tomcat**, and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.36 ALM-12057 Metadata Not Configured with the Task to Periodically Back Up Data to a Third-Party Server

Description

After the system is installed, it checks whether the task for periodically backing up metadata to the third-party server, and then performs the check hourly. If the task for periodically backing up metadata to a third-party server is not configured, a critical alarm is generated.

This alarm is cleared when a user creates such a backup task.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12057	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the metadata is not backed up to a third-party server, when the active/standby management nodes in the cluster are faulty and the local backup data is lost, if you want to use the backup package to restore the cluster metadata, no backup package is available.

Possible Causes

Metadata is not configured with the task to periodically back up data to a third-party server.

Procedure

Check whether periodic backup is configured for the metadata..

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On the FusionInsight Manager portal choose **O&M > Alarm > Alarms**.

Step 2 In the alarm list, click ⌄ in the row where the alarm is located and identify the data module from which the alarm is generated based on **Additional Information**.

Step 3 Choose **O&M** > **Backup and Restoration > Backup Management** > **Create**.

Step 4 Configure a backup task. The backup data to be configured is consistent with the data in Additional Information of the alarm.

Step 5 After the backup task is created successfully, wait for two minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information

Step 6 On FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 7 In the **Service** area, select **Controller** and click **OK**.

Step 8 Click ✎ in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.37 ALM-12061 Process Usage Exceeds the Threshold

Description

The system checks the usage of the omm process every 30 seconds. Users can run the **ps -o nlwp, pid, args, -u omm | awk '{sum+=$1} END {print "", sum}'** command to obtain the number of concurrent processes of user **omm**. Run the **ulimit -u** command to obtain the maximum number of processes that can be simultaneously opened by user **omm**.

Divide the number of concurrent processes by the maximum number to obtain the process usage of user **omm**. The process usage has a default threshold. This alarm is generated when the process usage exceeds the threshold.

If **Trigger Count** is **3** and the process usage is less than or equal to the threshold, this alarm is cleared. If **Trigger Count** is greater than **1** and the process usage is less than or equal to 90% of the threshold, this alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12061	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Service failure: When the process usage exceeds the threshold, you cannot switch to user **omm**. A new **omm** thread cannot be created. As a result, the job may fail to run.

Possible Causes

- The alarm threshold is improperly configured.
- The maximum number of processes (including threads) that can be concurrently opened by user **omm** is inappropriate.
- An excessive number of threads are opened at the same time.

Procedure

Check whether the alarm threshold or alarm hit number is properly configured.

Step 1 On the FusionInsight Manager, change the alarm threshold and **Trigger Count** based on the actual CPU usage.

Specifically, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Process** > **omm Process Usage** to change Trigger Count.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The alarm is generated when the process usage exceeds the threshold for the times specified by **Trigger Count**.

Set the alarm threshold based on the actual process usage. To check the process usage, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Process** > **omm Process Usage**.

Step 2 2 minutes later, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 3**.

Check whether the maximum number of processes (including threads) opened by user omm is appropriate.

Step 3 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and view the IP address of the host for which the alarm is generated.

Step 4 Log in to the host for which the alarm is generated as user **root**.

Step 5 Run the following command to check whether the maximum number of threads that can be concurrently opened by user **omm** is greater than or equal to 60000:

su - omm -c "ulimit -u"

- If it is, go to **Step 7**.
- If it is not, go to **Step 6**.

Step 6 Run the following command to change the maximum number of threads that can be concurrently opened by user **omm** to 60000:

vi /etc/security/limits.d/omm-nproc.conf

Run the following commands to switch to user **omm**, and restart the NodeAgent process. After two minutes, check whether the alarm is cleared.

su - omm

sh $BIGDATA_HOME/om-agent/nodeagent/bin/restart-agent.sh

- If it is, no further action is required.
- If it is not, go to **Step 11**.

Check whether an excessive number of processes are opened at the same time.

Step 7 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and view the IP address of the host for which the alarm is generated.

Step 8 Log in to the host where the alarm is generated as user **root**.

Step 9 Run the **ps -o nlwp, pid, lwp, args, -u omm|sort -n** command to check the numbers of threads used by the system.

The result is sorted based on the thread number. Analyze the top 5 thread numbers and check whether the threads are incorrectly used. (The PID column indicates the process ID.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, perform the following operations to stop the parent processes of the top 5 abnormally used processes without affecting services. Then, go to **Step 10**.

  a. Run the following command to query the parent process of the corresponding process:

  ps -ef | grep "*Process ID***"**

  The third column in the command output is the parent process ID.

  b. Run the following command to stop the parent process:

  kill -9 *parent process ID*

- If it is not, run the **ulimit -u** command to change the maximum number to be greater than 60000.

Step 10 Five minutes later, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager home page of the active clusters, choose **O&M** > **Log** > **Download**.

Step 12 Select **OmmServer** and **NodeAgent** from the **Service** and click **OK**.

Step 13 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.38 ALM-12062 OMS Parameter Configurations Mismatch with the Cluster Scale

Alarm Description

The system checks whether the OMS parameter configurations match with the cluster scale at each top hour. If the OMS parameter configurations do not meet the cluster scale requirements, the system generates this alarm. This alarm is automatically cleared when the OMS parameter configurations are modified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12062	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the name of the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the parameters configured for the current cluster are smaller than the configuration standard required by the cluster scale, problems such as job running delay and slow service page response may occur. In severe cases, the Agent or OMS process on the cluster node is abnormal. As a result, component jobs fail to be submitted and OMS data fails to be synchronized.

Possible Causes

The OMS parameter configurations mismatch with the cluster scale.

Handling Procedure

Check whether the OMS parameter configurations match with the cluster scale.

Step 1 Log in to the active management node of the MRS cluster as user **root**.

Run the following command to switch to user **omm**:

```
su - omm
```

Step 2 Run the following command to open the Manager log:

```
vi $BIGDATA_LOG_HOME/controller/scriptlog/modify_manager_param.log
```

Search for log "Current oms configurations can not support *xx* nodes", where *xx* indicates the number of nodes in the current cluster.

Step 3 Optimize the current cluster configuration by referring to **Optimizing Manager Configurations Based on the Number of Cluster Nodes**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 One hour later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 6 Select **Controller** from the **Service** and click **OK**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

Optimizing Manager Configurations Based on the Number of Cluster Nodes

Step 1 Log in to the active management node of the MRS cluster as user **root**.

Run the following command to switch to user **omm**:

```
su - omm
```

Step 2 Run the following command to switch the directory:

```
cd ${BIGDATA_HOME}/om-server/om/sbin
```

Step 3 Run the following command to view the current Manager configurations.

```
sh oms_config_info.sh -q
```

Step 4 Run the following command to specify the number of nodes in the current cluster:

Command format:

```
sh oms_config_info.sh -s Number of nodes
```

Example:

```
sh oms_config_info.sh -s 1000
```

Enter **y** as prompted.

```
The following configurations will be modified:
    Module      Parameter        Current             Target
    Controller  controller.Xmx   4096m          =>  16384m
    Controller  controller.Xms   1024m          =>  8192m       Controller
controller.node.heartbeat.error.threshold    30000                =>   60000
    Pms         pms.mem          8192m          =>  10240m
Do you really want to do this operation? (y/n):
```

The configurations are updated successfully if the following information is displayed:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
...
Operation has been completed. Now restarting OMS server.                [done]
Restarted oms server successfully.
```

NOTE

- OMS is automatically restarted during the configuration update process.
- Clusters with similar quantities of nodes have same Manager configurations. For example, when the number of nodes is changed from 100 to 101, no configuration item needs to be updated.

----End

9.11.39 ALM-12063 Unavailable Disk

Description

The system checks whether the data disk of the current host is available at the top of each hour. The system creates files, writes files, and deletes files in the mount directory of the disk. If the operations fail, the alarm is generated. If the operations succeed, the disk is available, and the alarm is cleared.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12063	Major	Yes

Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the name of the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DiskName	Specifies the disk for which the alarm is generated.

Impact on the System

Service failure: If you need to modify or use data on a disk that is unwritable or unreadable, the job may fail.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The permission of the disk mount directory is abnormal.
- There are disk bad sectors.

Procedure

Check whether the permission of the disk mount directory is normal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the IP address of the host and **DiskName** for the disk.

Step 2 Log in to the alarm host as user **root**.

Step 3 Run the **df -h |grep DiskName** command to obtain the mount point and check whether the permission of the mount directory is unwritable or unreadable.

- If it is, go to **Step 4**.
- If it is not, go to **Step 8**.

NOTE

If the permission of the mount directory is 000 or the owner is **root**, the mount directory is unreadable and unwritable.

Step 4 Modify the directory permission.

Step 5 One hour later, check whether this alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Step 6 Contact hardware engineers to rectify the disk.

Step 7 One hour later, check whether this alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 9 Select **NodeAgent** from the **Service** and click **OK**.

Step 10 Click ✎ in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.40 ALM-12064 Host Random Port Range Conflicts with Cluster Used Port

Alarm Description

The system checks whether the random port range of the host conflicts with the range of ports used by the Cluster system every hour. The alarm is generated if they conflict. The alarm is automatically cleared when the random port range of the host is changed to the normal range.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12064	Major	Yes

Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the name of the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- Authentication fails and jobs may fail to run.
- The process becomes faulty, which may affect the elastic scaling performance.
- If the Tomcat port is occupied, the login and query functions of FusionInsight Manager are affected.

Possible Causes

The random port range configuration is modified.

Procedure

Check the random port range of the system.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the IP address of the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Log in to the alarm host as user **root**.

Step 3 Run the **cat /proc/sys/net/ipv4/ip_local_port_range** command to obtain the random port range of the host and check whether the minimum value is smaller than 32768.

- If it is, go to **Step 4**.
- If it is not, go to **Step 7**.

Step 4 Run the **vim /etc/sysctl.conf** command to change the value of **net.ipv4.ip_local_port_range** to **32768 61000**. If this parameter does not exist, add the following configuration: **net.ipv4.ip_local_port_range = 32768 61000**.

Step 5 Run the **sysctl -p /etc/sysctl.conf** command for the modification to take effect.

Step 6 One hour later, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 8 Select **NodeAgent** for **Service** and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.41 ALM-12066 Trust Relationships Between Nodes Become Invalid

Alarm Description

The system checks whether the trust relationship between the active OMS node and other Agent nodes is normal every hour. The alarm is generated if the mutual trust fails. This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12066	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Some operations (such as restart, configuration synchronization, and instance status query) that need to connect to the node may fail. If the trust relationships between nodes are invalid, services may be affected.

Possible Causes

- The **/etc/ssh/sshd_config** configuration file on the cluster node is damaged.
- The password of user **omm** for logging in to the cluster node expired.

Handling Procedure

Check the status of the /etc/ssh/sshd_config configuration file.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm and click ⌄ to view the host list in the alarm details.

Step 2 Log in to the active OMS management node as user **root**.

Run the following command to switch to user **omm**:

```
su - omm
```

Step 3 Then run **ssh** to check whether the connection fails:

```
ssh host2
```

host2 indicates any other nodes except the OMS management node in the alarm details.

- If yes, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 6**.

Step 4 Open the **/etc/ssh/sshd_config** configuration file on the host specified by *host2* and check whether **AllowUsers** or **DenyUsers** includes the nodes specified by *host2*.

- If yes, go to **Step 5**.
- If no, contact OS experts.

Step 5 Modify the whitelist or blacklist to ensure that user **omm** is in the whitelist or not in the blacklist. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check the status of the password of user omm.

Step 6 Check the interaction information of the **ssh** command.

- If the password of user **omm** is required, go to **Step 7**.
- If message "Enter passphrase for key '/home/omm/.ssh/id_rsa':" is displayed, go to **Step 9**.
- If the connection is set up successfully, run the following command to restart **ssh-agent** and then go to **Step 3**:

```
ps -ef|grep ssh-agent |grep -v grep |awk '{print $2}' | xargs kill -9
```

Step 7 Check the trust list (**/home/omm/.ssh/authorized_keys**) of user **omm** on the OMS node and host2 node. Check whether the trust list contains the public key file (**/home/omm/.ssh/id_rsa.pub**) of user **omm** on the peer host.

- If yes, contact OS experts.
- If no, add the public key of user **omm** of the peer host to the trust list of the local host.

Step 8 Add the public key of user **omm** of the peer host to the trust list of the local host. Run the **ssh** command, for example, **ssh host2**, on each node in the alarm details to check whether the connection fails. (***host2*** is a node other than the OMS node in the alarm details.)

- If yes, go to **Step 9**.
- If no, check whether the alarm is cleared. If the alarm is cleared, no further action is required; otherwise, go to **Step 9**.

Collect the fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Select **Controller** for **Service** and click **OK**.

Step 11 Click  in the upper right corner to set the log collection time range. Generally, the time range is 10 minutes before and after the alarm generation time. Click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

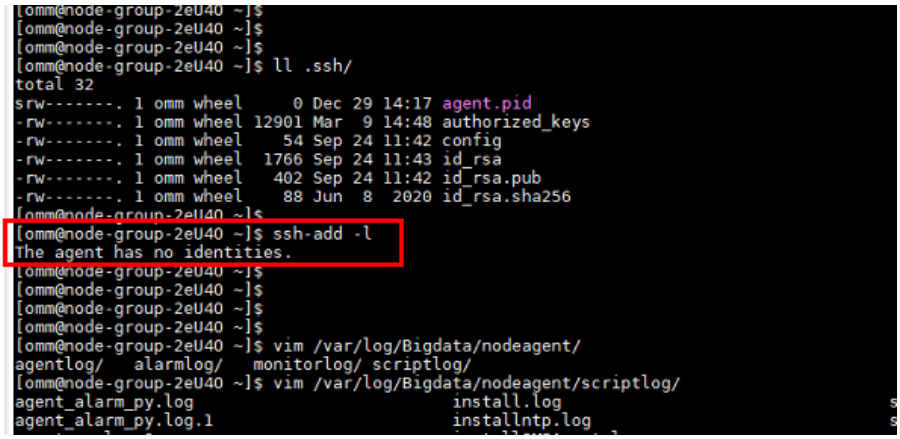
Related Information

Perform the following steps to handle abnormal trust relationships between nodes:

NOTICE

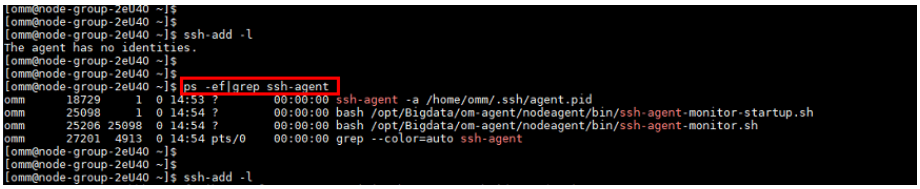
- Perform this operation as user **omm**.
- If the network between nodes is disconnected, rectify the network fault first. Check whether the two nodes are connected to the same security group and whether **hosts.deny** and **hosts.allow** are set.

1. Run the **ssh-add -l** command on both nodes to check whether any identities exist.



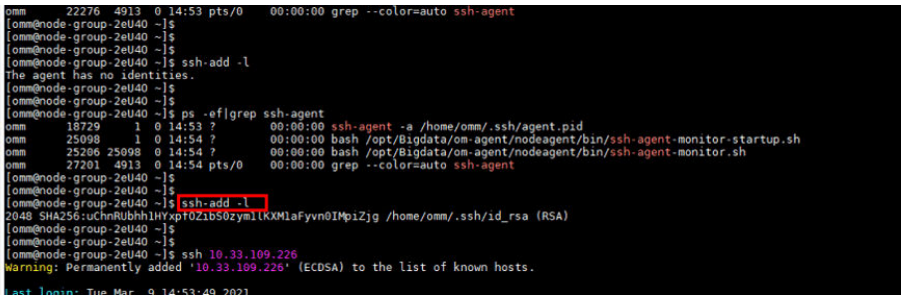
   - If yes, go to 4.
   - If no, go to 2.

2. If no identities are displayed, run the **ps -ef|grep ssh-agent** command to find the **ssh-agent** process, stop the process, and wait for the process to automatically restart.



3. Run the **ssh-add -l** command to check whether the identities have been added. If yes, manually run the **ssh** command to check whether the trust relationship is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



4. If identities exist, check whether the **/home/omm/.ssh/authorized_keys** file contains the information in the **/home/omm/.ssh/id_rsa.pub** file of the peer node. If it does not, manually add the information.
5. Check whether the permissions on the files in the **/home/omm/.ssh** directory are modified.
6. Check the **/var/log/Bigdata/nodeagent/scriptlog/ssh-agent-monitor.log** file.
7. If the **/home** directory of user **omm** is deleted, contact MRS support personnel for assistance.

9.11.42 ALM-12067 Tomcat Resource Is Abnormal

Alarm Description

HA checks the Tomcat resources of Manager every 85 seconds. This alarm is generated when HA detects that the Tomcat resources are abnormal for two consecutive times.

This alarm is cleared when HA detects that the Tomcat resources become normal.

Resource Type of Tomcat is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new Tomcat resources have been enabled on the new active Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12067	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

An active/standby Manager switchover may occur. The Manager and component web UIs are unavailable. The cluster management function cannot be provided for upper-layer web applications, and users may fail to log in to the Manager and component web UIs.

Possible Causes

- The Tomcat directory permission is abnormal, and the Tomcat process is abnormal.

Handling Procedure

Check whether the permission on the Tomcat directory is normal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the IP address of the host for which the alarm is generated.

Step 2 Log in to the alarm host as user **root**.

Step 3 Run the **su - omm** command to switch to user **omm**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/tomcat.log** command to check whether the Tomcat resource log contains keyword **Cannot find *XXX*** and rectify the file permission based on the keyword.

Step 5 After 5 minutes, check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 In the **Services** area, select **OmmServer** and **Tomcat**, and click **OK**.

Step 8 Click ✎ in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.43 ALM-12068 ACS Resource Exception

Alarm Description

HA checks the ACS resources of Manager every 80 seconds. This alarm is generated when HA detects that the ACS resources are abnormal for two consecutive times.

This alarm is cleared when HA detects that the ACS resources are normal.

Resource Type of ACS is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new ACS resources have been enabled on the new active Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12068	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

An active/standby Manager switchover may occur. The security authentication and user management functions cannot be provided for ACS upper-layer applications. As a result, you may fail to log in to Manager and component web UIs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The ACS process is abnormal.

Handling Procedure

Check whether the ACS process is normal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the alarm host as user **root**.

Step 3 Run the **su - omm** command and then **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** to check whether the status of the ACS resources managed by the HA is normal. In the single-node system, the ACS resource is in the normal state. In the dual-node system, the ACS resource is in the normal state on the active node and in the stopped state on the standby node.

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/acs.log** command to check whether the ACS resource log of HA contains the keyword **ERROR**. If yes, analyze the logs to locate the resource exception cause and fix the exception.

Step 5 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 In the **Services** area, select **Controller** and **OmmServer**, and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.44 ALM-12069 AOS Resource Exception

Alarm Description

HA checks the AOS resources of Manager every 81 seconds. This alarm is generated when HA detects that the AOS resources are abnormal for two consecutive times.

This alarm is cleared when HA detects that the AOS resources become normal.

Resource Type of AOS is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new AOS resources have been enabled on the new active Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12069	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

An active/standby Manager switchover may occur. Tenant and role management cannot be provided for AOS upper-layer applications. As a result, you may fail to log in to Manager and component web UIs.

Possible Causes

The AOS process is abnormal.

Handling Procedure

Check whether the AOS process is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the alarm host as user **root**.

Step 3 Run the **su - omm** command and then **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** to check whether the status of the AOS resources managed by the HA is normal. In the single-node system, the AOS resource is in the normal state. In the dual-node system, the AOS resource is in the normal state on the active node and in the stopped state on the standby node.

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/aos.log** command to check whether the AOS resource log of HA contains the keyword **ERROR**. If yes, analyze the logs to locate the resource exception cause and fix the exception.

Step 5 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 In the **Services** area, select **Controller** and **OmmServer**, and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.45 ALM-12070 Controller Resource Is Abnormal

Alarm Description

HA checks the controller resources of Manager every 80 seconds. This alarm is generated when HA detects that the controller resources are abnormal for 2 consecutive times.

This alarm is cleared when the Controller resource is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Resource Type of Controller is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new Controller resources have been enabled on the new active FusionInsight Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12070	Major	Yes

Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the name of the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The alarm persists for a long time, causing frequent active/standby switchovers of FusionInsight Manager. As a result, users cannot log in to FusionInsight Manager and perform O&M operations.
- The Controller process repeatedly restarts, which may cause the native UI of the service login failure.

Possible Causes

The Controller process is abnormal.

Procedure

Check whether the controller process is normal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command to switch to user **omm**.Run the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the Controller resources

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

managed by the HA is normal. In the single-node system, the Controller resource is in the normal state. In the dual-node system, the Controller resource is in the normal state on the active node and in the stopped state on the standby node.

- If it is, go to **Step 6**.
- If it is not, go to **Step 4**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/controller.log** command to view the Controller resource logs, and run the **vi $BIGDATA_LOG_HOME/controller/controller.log** command to view the Controller running logs, check whether the keyword **ERROR** exists. Analyze the logs to locate and rectify the fault.

Step 5 Five minutes later, check whether this alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Select **Controller** and **OmmServer** for **Service** and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour before and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.46 ALM-12071 Httpd Resource Is Abnormal

Description

HA checks the httpd resources of Manager every 120 seconds. This alarm is generated when HA detects that the httpd resources are abnormal for 10 consecutive times.

This alarm is cleared when the httpd resource is normal.

Resource Type of httpd is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new httpd resources have been enabled on the new active FusionInsight Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12071	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The alarm persists for a long time, causing frequent active/standby switchovers of FusionInsight Manager. As a result, users cannot log in to FusionInsight Manager and perform O&M operations.
- The httpd process is repeatedly restarts, which may lead to the failure to visit the native service UI.

Possible Causes

The httpd process is abnormal.

Procedure

Check whether the httpd process is abnormal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command to switch to user **omm**.

Step 4 Run the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the httpd resources managed by the HA is normal. In the single-node system, the httpd resource is in the normal state. In the dual-node system, the httpd resource is in the normal state on the active node and in the stopped state on the standby node.

- If it is, go to **Step 7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is not, go to **Step 5**.

Step 5 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/httpd.log** command to view the httpd resource logs, check whether the keyword **ERROR** exists. Analyze the logs to locate and rectify the fault.

Step 6 Five minutes later, check whether this alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 8 Select **Controller** and **OmmServer** for **Service** and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.47 ALM-12072 FloatIP Resource Is Abnormal

Description

HA checks the floatip resources of Manager every 9 seconds. This alarm is generated when HA detects that the floatip resources are abnormal for 3 consecutive times.

This alarm is cleared when the FloatIP resource is normal.

Resource Type of FloatIP is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new FloatIP resources have been enabled on the new active FusionInsight Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12072	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The alarm persists for a long time, causing frequent active/standby switchovers of FusionInsight Manager. As a result, users cannot log in to FusionInsight Manager and perform O&M operations.
- The FloatIP process is repeatedly restarts, which may lead to the failure to visit the native service UI.

Possible Causes

The floating IP address is abnormal.

Procedure

Check the floating IP address status of the active management node.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm and click ⌄ to view the host address and resource name for which the alarm is generated.

Step 2 Log in to the active management node as user **root**.

Step 3 Run the following command, go to the **${BIGDATA_HOME}/om-server/om/sbin/** directory.

su - omm

cd ${BIGDATA_HOME}/om-server/om/sbin/

Step 4 Run the **sh status-oms.sh** command, and execute the **status-oms.sh** script to check whether the floating IP address of the active FusionInsight Manager is normal. View the command output, locate the row where **ResName** is **floatip**, and check whether the following information is displayed.

For example:

```
10-10-10-160   floatip  Normal  Normal   Single_active
```

- If it is, go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is not, go to **Step 5**.

Step 5 Run the **ifconfig** command to check whether the NIC with the floating IP address exists.

- If it does, go to **Step 8**.
- If it does not, go to **Step 6**.

Step 6 Run the **ifconfig** *NIC name Floating IP address* netmask *Subnet mask* command to reconfigure the NIC with the floating IP address. (For example, **ifconfig eth0 10.10.10.102 netmask 255.255.255.0**).

Step 7 Five minutes later, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 9 Select **Controller** and **OmmServer** for **Service** and click **OK**.

Step 10 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 1 hour before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 11 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.48 ALM-12073 CEP Resource Is Abnormal

Description

HA checks the cep resources of Manager every 60 seconds. This alarm is generated when HA detects that the cep resources are abnormal for 2 consecutive times.

This alarm is cleared when the CEP resource is normal.

Resource Type of CEP is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new CEP resources have been enabled on the new active FusionInsight Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12073	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The alarm persists for a long time, causing frequent active/standby switchovers of FusionInsight Manager. As a result, users cannot log in to FusionInsight Manager and perform O&M operations.
- The CEP process repeatedly restarts. As a result, monitoring data collection fails during the alarm reporting period. In severe cases, monitoring data during the alarm reporting period may be lost.

Possible Causes

The CEP process is abnormal.

Procedure

Check whether the CEP process is abnormal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command and then the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the CEP resources managed by the HA is normal. In the single-node system, the CEP resource is in the normal state. In the dual-node system, the CEP resource is in the normal state on the active node and in the stopped state on the standby node.

- If it is, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is not, go to **Step 4**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/cep/cep.log** and **vi $BIGDATA_LOG_HOME/omm/oms/cep/scriptlog/cep_ha.log** commands to view the CEP resource logs, check whether the keyword **ERROR** exists. Analyze the logs to locate and rectify the fault.

Step 5 Five minutes later, check whether this alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Select **Controller** and **OmmServer** for **Service** and click **OK**.

Step 8 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 1 hour before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.49 ALM-12074 FMS Resource Is Abnormal

Description

HA checks the fms resources of Manager every 60 seconds. This alarm is generated when HA detects that the fms resources are abnormal for 2 consecutive times.

This alarm is cleared when the FMS resource is normal.

Resource Type of FMS is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new FMS resources have been enabled on the new active FusionInsight Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12074	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The alarm persists for a long time, causing frequent active/standby switchovers of FusionInsight Manager. As a result, users cannot log in to FusionInsight Manager and perform O&M operations.
- The FMS process repeatedly restarts. As a result, the alarm data reported during the alarm reporting period is abnormal. In severe cases, the alarm data reported during the alarm reporting period may fail to be reported and cleared.

Possible Causes

The FMS process is abnormal.

Procedure

Check whether the FMS process is abnormal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command and then the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the FMS resources managed by the HA is normal. In the single-node system, the FMS resource is in the normal state. In the dual-node system, the FMS resource is in the normal state on the active node and in the stopped state on the standby node.

- If it is, go to **Step 6**.
- If it is not, go to **Step 4**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/fms/fms.log** and **vi $BIGDATA_LOG_HOME/omm/oms/fms/scriptlog/fms_ha.log** commands to view the FMS resource logs, check whether the keyword **ERROR** exists. Analyze the logs to locate and rectify the fault.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 5 minutes later, check whether this alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Select **Controller** and **OmmServer** for **Service** and click **OK**.

Step 8 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 1 hour before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.50 ALM-12075 PMS Resource Is Abnormal

Description

HA checks the pms resources of Manager every 55 seconds. This alarm is generated when HA detects that the pms resources are abnormal for three consecutive times.

This alarm is cleared when the PMS resource is normal.

Resource Type of PMS is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new PMS resources have been enabled on the new active FusionInsight Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12075	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The alarm persists for a long time, causing frequent active/standby switchovers of FusionInsight Manager. As a result, users cannot log in to FusionInsight Manager and perform O&M operations.
- The PMS process repeatedly restarts. As a result, monitoring data collection fails during the alarm reporting period. In severe cases, monitoring data may be lost during the alarm reporting period.

Possible Causes

The PMS process is abnormal.

Procedure

Check whether the PMS process is abnormal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command and then the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the PMS resources managed by the HA is normal. In the single-node system, the PMS resource is in the normal state. In the dual-node system, the PMS resource is in the normal state on the active node and in the stopped state on the standby node.

- If it is, go to **Step 6**.
- If it is not, go to **Step 4**.

Step 4 Run the **vi $BIGDATA_LOG_HOME/omm/oms/pms/pms.log** and **vi $BIGDATA_LOG_HOME/omm/oms/pms/scriptlog/pms_ha.log** commands to view the PMS resource logs, check whether the keyword **ERROR** exists. Analyze the logs to locate and rectify the fault.

Step 5 Five minutes later, check whether this alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Select **Controller** and **OmmServer** for **Service** and click **OK**.

Step 8 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 1 hour before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.51 ALM-12076 GaussDB Resource Is Abnormal

Description

HA checks the Manager database every 10 seconds. This alarm is generated when HA detects that the database is abnormal for 3 consecutive times.

This alarm is cleared when the database is normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12076	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the database is abnormal, all core services and related service processes of Manager, such as the alarm, monitoring, and query functions, are affected.

Possible Causes

An exception occurs in the database.

Procedure

Check the database status of the active and standby management nodes.

Step 1 Log in to the active and standby management nodes as user **root**. . Run the **su - ommdba** command to switch to user **ommdba**, and run the **gs_ctl query** command. Check whether the following information is displayed.

Command output of the active management node:

```
Ha state:
        LOCAL_ROLE                  : Primary
        STATIC_CONNECTIONS          : 1
        DB_STATE                    : Normal
        DETAIL_INFORMATION          : user/password invalid
Senders info:
        No information
Receiver info:
        No information
```

Command output of the standby management node:

```
Ha state:
        LOCAL_ROLE                  : Standby
        STATIC_CONNECTIONS          : 1
        DB_STATE                    : Normal
        DETAIL_INFORMATION          : user/password invalid
Senders info:
        No information
Receiver info:
        No information
```

- If it is, go to **Step 3**.
- If it is not, go to **Step 2**.

Step 2 Contact the network administrator to check whether the network is faulty.

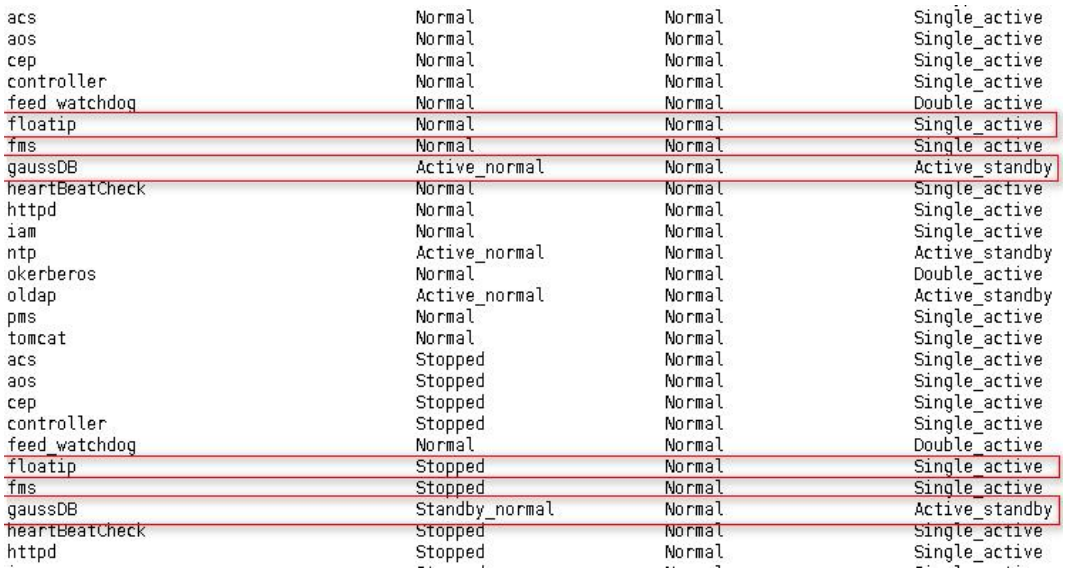
- If it is, go to **Step 3**.
- If it is not, go to **Step 5**.

Step 3 Five minutes later, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Log in to the active and standby management nodes, run the **su - omm** command to switch to user **omm**, go to the **${BIGDATA_HOME} /om-server/om/sbin/** directory, and run the **status-oms.sh** script to check whether the floating IP addresses and GaussDB resources of the active and standby FusionInsight Managers are in the status shown in the following figure.



acs	Normal	Normal	Single_active
aos	Normal	Normal	Single_active
cep	Normal	Normal	Single_active
controller	Normal	Normal	Single_active
feed_watchdog	Normal	Normal	Double_active
floatip	Normal	Normal	Single_active
fms	Normal	Normal	Single_active
gaussDB	Active_normal	Normal	Active_standby
heartBeatCheck	Normal	Normal	Single_active
httpd	Normal	Normal	Single_active
iam	Normal	Normal	Single_active
ntp	Active_normal	Normal	Active_standby
okerberos	Normal	Normal	Double_active
oldap	Active_normal	Normal	Active_standby
pms	Normal	Normal	Single_active
tomcat	Normal	Normal	Single_active
acs	Stopped	Normal	Single_active
aos	Stopped	Normal	Single_active
cep	Stopped	Normal	Single_active
controller	Stopped	Normal	Single_active
feed_watchdog	Normal	Normal	Double_active
floatip	Stopped	Normal	Single_active
fms	Stopped	Normal	Single_active
gaussDB	Standby_normal	Normal	Active_standby
heartBeatCheck	Stopped	Normal	Single_active
httpd	Stopped	Normal	Single_active

- If they are, find the alarm in the alarm list and manually clear the alarm.
- If they are not, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 6 Select **OmmServer** for **Service** and click **OK**.

Step 7 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 8 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.52 ALM-12077 User omm Expired

Description

The system starts at 00:00 every day to check whether user **omm** has expired every eight hours. This alarm is generated if the user account has expired.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the expiration time of user **omm** is changed and the user account status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12077	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

User **omm** has expired. The node trust relationship is unavailable, and FusionInsight Manager cannot manage the services.

Possible Causes

User **omm** has expired.

Procedure

Check whether user omm in the system has expired.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l omm** command to view the information about the password of user **omm**.

Step 2 View the value of **Account expires** to check whether the user configurations have expired.

NOTE

If the parameter value is **never**, the user configurations never expire.

- If they do, go to **Step 3**.
- If they do not, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Run the **chage -E 'yyyy-MM-dd' omm** command to set the expiration time of user **omm**. Eight hours later, check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.53 ALM-12078 Password of User omm Expired

Description

The system starts at 00:00 every day to check whether the password of user **omm** has expired every 8 hours. This alarm is generated if the password has expired.

This alarm is cleared when the expiration time of user **omm** password is changed and the user password status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12078	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The password of user **omm** has expired. The node trust relationship is unavailable, and FusionInsight Manager cannot manage the services. The **crontab** scheduled task cannot be executed, affecting the ClickHouse service.

Possible Causes

The password of user **omm** has expired.

Procedure

Check whether the password of user omm in the system has expired.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l omm** command to view the information about the password of user **omm**.

Step 2 View the value of **Password expires** to check whether the user configurations have expired.

NOTE

If the parameter value is **never**, the user configurations never expire.

- If they do, go to **Step 3**.
- If they do not, go to **Step 4**.

Step 3 Run the **chage -M** '*days*' **omm** command to set the validity period of the password for user **omm**. Eight hours later, check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.54 ALM-12079 User omm Is About to Expire

Description

The system starts at 00:00 every day to check whether user **omm** is about to expire every 8 hours. This alarm is generated if the user account will expire no less than 15 days later.

This alarm is cleared when the expiration time of user **omm** is changed and the user account status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12079	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

User **omm** has expired. The node trust relationship is unavailable, and FusionInsight Manager cannot manage the services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The account of user **omm** is about to expire.

Procedure

Check whether user omm is about to expire.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l omm** command to view the information about the password of user **omm**.

Step 2 View the value of **Account expires** to check whether the user configurations are about to expire.

NOTE

If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password are about to expire within 15 days.

- If they are, go to **Step 3**.
- If they are not, go to **Step 4**.

Step 3 Run the **chage -E** *'yyyy-MM-dd'* **omm** command to set the validity period of user **omm**. Eight hours later, check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.55 ALM-12080 Password of User omm Is About to Expire

Description

The system starts at 00:00 every day to check whether the password of user **omm** is about to expire every 8 hours. This alarm is generated if the password will expire no less than 15 days later.

This alarm is cleared when the expiration time of user **omm** password is reset and the user password status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12080	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The password of user **omm** has expired. The node trust relationship is unavailable, and FusionInsight Manager cannot manage the services. The **crontab** scheduled task cannot be executed, affecting the ClickHouse service.

Possible Causes

The password of user **omm** is about to expire.

Procedure

Check whether the password of user omm in the system is about to expire.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l omm**command to view the information about the password of user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 View the value of **Password expires** to check whether the user configurations are about to expire.

NOTE

If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password are about to expire within 15 days.

- If they are, go to **Step 3**.
- If they are not, go to **Step 4**.

Step 3 Run the **chage -M '***days***' omm** command to set the validity period of the password for user **omm**. Eight hours later, check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.56 ALM-12081 User ommdba Expired

Description

The system starts at 00:00 every day to check whether user **ommdba** has expired every 8 hours. This alarm is generated if the user account has expired.

This alarm is cleared when the expiration time of user **ommdba** is reset and the user account status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12081	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The OMS database cannot be managed and data cannot be accessed.

Possible Causes

The account of user **ommdba** for the host has expired.

Procedure

Check whether user ommdba has expired.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l ommdba** command to view the information about the password of user **ommdba**.

Step 2 View the value of **Account expires** to check whether the user configurations have expired.

NOTE

If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password have expired.

- If they do, go to **Step 3**.
- If they do not, go to **Step 4**.

Step 3 Run the **chage -E** *'yyyy-MM-dd'* **omm** command to set the validity period of user **ommdba**. Eight hours later, check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.57 ALM-12082 User ommdba Is About to Expire

Description

The system starts at 00:00 every day to check whether user **ommdba** is about to expire every 8 hours. This alarm is generated if the user account will expire no less than 15 days later.

This alarm is cleared when the expiration time of user **ommdba** is reset and the user account status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12082	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The OMS database cannot be managed and data cannot be accessed.

Possible Causes

The account of user **ommdba** for the host is about to expire.

Procedure

Check whether user ommdba is about to expire.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l ommdba** command to view the information about user **ommdba**.

Step 2 View the value of **Account expires** to check whether the user configurations are about to expire.

NOTE

If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password are about to expire within 15 days.

- If they are, go to **Step 3**.
- If they are not, go to **Step 4**.

Step 3 Run the **chage -E** *'yyyy-MM-dd'* **ommdba** command to set the validity period of user **ommdba**. Eight hours later, check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click  in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.58 ALM-12083 Password of User ommdba Is About to Expire

Description

The system starts at 00:00 every day to check whether the password of user **ommdba** is about to expire every 8 hours. This alarm is generated if the password is about to expire no less than 15 days later.

This alarm is cleared when the expiration time of user **ommdba** password is reset and the user password status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12083	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The OMS database cannot be managed and data cannot be accessed.

Possible Causes

The password of user **ommdba** is about to expire.

Procedure

Check whether the password of user ommdba in the system is about to expire.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l ommdba** command to view the information about the password of user **ommdba**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 View the value of **Password expires** to check whether the user configurations are about to expire.

NOTE

If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password are about to expire within 15 days.

- If they are, go to **Step 3**.
- If they are not, go to **Step 4**.

Step 3 Run the **chage -M** *'days'* **ommdba** command to set the validity period of the password for user **ommdba**. Eight hours later, check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.59 ALM-12084 Password of User ommdba Expired

Description

The system starts at 00:00 every day to check whether the password of user **ommdba** has expired every 8 hours. This alarm is generated if the password has expired.

This alarm is cleared when the expiration time of user **ommdba** password is reset and the user password status becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12084	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The password of user **ommdba** has expired. The node trust relationship is unavailable, and FusionInsight Manager cannot manage the services.

Possible Causes

The password of user **ommdba** for the host has expired.

Procedure

Check whether the password of user ommdba in the system has expired.

Step 1 Log in to the faulty node in the cluster as user **root**.

Run the **chage -l ommdba** command to view the information about the password of user **ommdba**.

Step 2 View the value of **Password expires** to check whether the user configurations have expired.

NOTE

If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password have expired.

- If they do, go to **Step 3**.
- If they do not, go to **Step 4**.

Step 3 Run the **chage -M** *'days'* **ommdba** command to set the validity period of the password for user **ommdba**. Eight hours later, check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Select **NodeAgent** for **Service** and click **OK**.

Step 6 Click  in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.60 ALM-12085 Service Audit Log Dump Failure

Description

The system dumps service audit logs at 03:00 every day and stores them on the OMS node. This alarm is generated when the dump fails. This alarm is cleared when the next dump succeeds.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12085	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the audit logs of a component fail to be dumped, the audit logs cannot be retrieved if they are aged locally. This affects service analysis and troubleshooting of the component.

Possible Causes

- The service audit logs are oversized.
- The OMS backup storage space is insufficient.
- The storage space of a host where the service is located is insufficient.

Procedure

Check whether the service audit logs are oversized.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the IP address of the host for which the alarm is generated and additional information.

Step 2 Log in to the host where the alarm is generated as user **root**.

Step 3 Run the **vi ${BIGDATA_LOG_HOME}/controller/scriptlog/getLogs.log** command to search for the keyword "LOG SIZE is more than 5000MB", which indicates whether the keyword can be found.

- If it can, go to **Step 4**.
- If it cannot, go to **Step 5**.

Step 4 Check whether the oversized service audit logs are caused by exceptions.

The OMS backup storage space is insufficient.

Step 5 Run the **vi ${BIGDATA_LOG_HOME}/controller/scriptlog/getLogs.log** command to search for the keyword "Collect log failed, too many logs on", which indicates whether the keyword can be found.

- If it can, obtain the host IP address following the keyword "Collect log failed, too many logs on", and go to **Step 6**.
- If it cannot, go to **Step 11**.

Step 6 Log in to the host with the IP address obtained in **Step 5** as user **root**.

Step 7 Run the **vi ${BIGDATA_LOG_HOME}/nodeagent/scriptlog/collectLog.log** command to check whether the keyword "log size exceeds" can be searched.

- If it can, go to **Step 9**.
- If it cannot, go to **Step 8**.

Step 8 Check whether the alarm additional information contains the keyword "no enough space".

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Perform the following operations to expand the disk capacity (only for MRS 3.1.2 and earlier versions) or reduce the maximum number of audit log backups:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Expand the capacity of the OMS node.
- Run the following command to edit the file and decrease the value of **MAX_NUM_BK_AUDITLOG**.

  vi ${CONTROLLER_HOME}/etc/om/componentsauditlog.properties

Step 10 In the next execution period, 03:00, check whether the alarm is cleared.

- If yes, no further action is required.
- If it is not, go to **Step 11**.

Check whether the space of the host where the service is located is insufficient.

Step 11 Run the **vi ${BIGDATA_LOG_HOME}/controller/scriptlog/getLogs.log** command to search for the keyword "Collect log failed, no enough space on *hostIp* ", which indicates whether the keyword can be found.

- If it can, obtain the IP address of the abnormal host and go to **Step 12**.
- If it cannot, go to **Step 15**.

Step 12 Log in to the host with the IP address obtained as user **root**, and run the **df "$BIGDATA_HOME/tmp" -lP | tail -1 | awk '{print ($4/1024)}'** command to obtain the remaining space of the host log directory. Check whether the value is less than 1000 MB.

- If it is, go to **Step 13**.
- If it is not, go to **Step 15**.

Step 13 Expand the capacity of the node

Step 14 In the next execution period, 03:00, check whether the alarm is cleared.

- If yes, no further action is required.
- If it is not, go to **Step 15**.

Collect fault information.

Step 15 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 16 Select **Controller** for **Service** and click **OK**.

Step 17 Click  in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 18 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.61 ALM-12087 System Is in the Upgrade Observation Period

Description

The system checks whether it is in the upgrade observation period at 00:00 every day and checks whether the duration that it has been in the upgrade observation state exceeds the preset upgrade observation period, 10 days by default. This alarm is generated when the system is in the upgrade observation period and the duration that the system has been in the upgrade observation state exceeds the preset period (10 days by default). This alarm is automatically cleared if the system exits the upgrade observation period after the user performs a rollback or submission.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12087	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Upgrade Observation Period (Days)	Specifies the days that the system is in the upgrade observation period.

Impact on the System

During the upgrade observation period, do not add or delete users, instances, roles, services, hosts, or resource pools that affect the management topology.

Possible Causes

The upgrade task is not submitted a specified period of time (10 days by default) after the system upgrade.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check whether the system is in the upgrade observation period.

Step 1 Log in to the active management node as user **root**.

Step 2 Run the following commands to switch to user **omm** and log in to the **omm** database:

su - omm

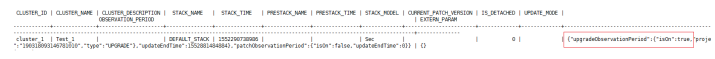
gsql -U omm -W *omm database password* **-p 20015**

Step 3 Run the **select * from OM_CLUSTERS** command to view cluster information.

Step 4 Check whether the value of **upgradObservationPeriod isON** is **true**, as shown in **Figure 9-6**.

- If it is, the system is in the upgrade observation period. Use the UpdateTool to submit the upgrade task. For details, see the upgrade guide of the corresponding version.
- If it is not, go to **Step 6**.

Figure 9-6 Cluster information



Step 5 In the early morning of the next day, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Controller** from the **Service** and click **OK**.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.62 ALM-12089 Inter-Node Network Is Abnormal

Alarm Description

The alarm module checks the network health status of nodes in the cluster every 10 seconds. This alarm is generated when the network between two nodes is unreachable or the network status is unstable.

This alarm is cleared when the network recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12089	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- Data transmission becomes slow or interrupted. Data may be lost or incomplete.
- Task scheduling is affected. For example, Yarn tasks cannot be executed properly or fail to be executed due to timeout.
- Data processing is affected. For example, HDFS data synchronization fails or the data is inaccurate.
- System performance deteriorates. The efficiency and quality of data processing is low.

Possible Causes

- The node breaks down.
- The network is faulty.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the network health status.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, click ⌄ to view the description in additional information. Record the source IP address and destination IP address of the node for which the alarm is reported.

Step 2 Log in to the node that the alarm is generated for. On the node, run **ping** the destination node to check whether the network between the two nodes is normal.

```
ping IP address of the destination node
```

- If yes, go to **6**.
- If no, go to **3**.

Check the node status.

Step 3 On FusionInsight Manager, click **Host** and check whether the host list contains the faulty node to determine whether the faulty node has been removed from the cluster.

- If yes, go to **5**.
- If no, go to **4**.

Step 4 Check whether the faulty node is powered off.

- If yes, start the faulty node and go to **Step 2**.
- If no, contact related personnel to find root cause. If you need to remove the faulty nodes from the cluster, go to **Step 5**. Otherwise, go to **Step 6**.

Step 5 Remove the faulty node from the **$NODE_AGENT_HOME/etc/agent/hosts.ini** file on all nodes in the cluster, clear the **/var/log/Bigdata/unreachable/unreachable_ip_info.log** file, and clear the alarm.

Step 6 Wait 30 seconds and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **OmmAgent** from the **Service** and click **OK**.

Step 9 Click ✎ in the upper right corner to set the log collection time range. Generally, the time range is 10 seconds before and after the alarm generation time. Click **Download**.

Step 10 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.63 ALM-12091 Abnormal disaster Resources

Alarm Description

HA checks the disaster resources of Manager every 86 seconds. This alarm is generated when HA detects that the disaster resources have been abnormal for 10 consecutive times.

This alarm is cleared when HA detects that the disaster resources become normal.

Resource Type of disaster is **Single-active**. Active/Standby switchover will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new disaster resources have been enabled on the new active Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12091	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The active/standby Manager switchover occurs.
- The disaster process restarts repeatedly, which may cause active/standby DR to be unavailable.

Possible Causes

The disaster process is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check whether the disaster process is normal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command to switch to user **omm**.

Step 4 Run the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the disaster resources managed by the HA is normal. In the single-node system, the disaster resource is in the normal state. In the dual-node system, the disaster resource is in the normal state on the active node and in the stopped state on the standby node.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

Step 5 Run the **vi ${BIGDATA_LOG_HOME}/disaster/disaster.log** command to check whether the disaster resource log of HA contains the keyword **ERROR**. If yes, analyze the logs to locate the resource exception cause and fix the exception.

Step 6 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, select **Disaster** for the target cluster, and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.64 ALM-12101 AZ Unhealthy

Description

After the AZ DR function is enabled, the system checks the AZ health status every 5 minutes. This alarm is generated when the system detects that the AZ is subhealthy or unhealthy. This alarm is cleared when the AZ becomes healthy.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12101	Critical	Yes

Parameters

Parameter	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
AZName	Specifies the AZ for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The health status of an AZ is determined by whether the health status of storage resources (HDFS), computing resources (Yarn), and key roles in the AZ exceeds the configured threshold.

An AZ is subhealthy when:

- The computing resources (Yarn) are unhealthy, but the storage resources (HDFS) are healthy. Tasks cannot be submitted to the local AZ, but data can still be read and written in the local AZ.
- The computing resources (Yarn) are healthy, but some storage resources (HDFS) are unhealthy. Tasks can be submitted to the local AZ, and some data can be read and written in the local AZ. This depends on the locality of data detected by Spark/Hive scheduling.

An AZ is unhealthy when:

- The computing resources (Yarn) are healthy, but the storage resources (HDFS) are unhealthy. Although tasks can be submitted to the local AZ, data cannot be read or written in the local AZ. As a result, the tasks submitted to the local AZ are invalid.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The computing resources (Yarn) and storage resources (HDFS) are unhealthy. Tasks cannot be submitted to the local AZ, and data cannot be read or written in the local AZ.
- The health status of key roles except Yarn and HDFS is lower than the configured threshold.

Possible Causes

- The computing resources (Yarn) are unhealthy.
- The storage resources (HDFS) are unhealthy.
- Some storage resources (HDFS) are unhealthy.
- Key roles except Yarn and HDFS are unhealthy.

Procedure

Disable the DR drill.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Cross-AZ HA**. The Cross-AZ HA page is displayed.

Step 2 In the AZ DR list, check whether **Perform DR Drill** in the **Operation** column of the AZ whose health status is **Unhealthy** is gray.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 Click **Restore** in the **Operation** column of the target AZ. Wait 2 minutes and refresh the page to view the health status of the AZ. Check whether the health status is normal.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect the fault information.

Step 4 Log in to the active management node as user **root**.

Step 5 View logs of unhealthy services.

- HDFS log files are stored in **/var/log/Bigdata/hdfs/nn/hdfs-az-state.log**.
- Yarn log files are stored in **/var/log/Bigdata/yarn/rm/yarn-az-state.log**.
- For other services, view the service health check logs in the corresponding service log directory.

Step 6 Contact O&M personnel and provide detailed log file information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.65 ALM-12102 AZ HA Component Is Not Deployed Based on DR Requirements

Alarm Description

The alarm module checks the deployment status of AZ HA components every 5 minutes. This alarm is generated when the components that support DR are not deployed based on DR requirements after AZ is enabled. This alarm is cleared when the components are deployed based on DR requirements.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12102	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

Impact on the System

The cross-AZ HA capability of a single cluster is affected.

Possible Causes

The roles of the components that support DR are not deployed based on DR requirements.

Handling Procedure

Obtain alarm information.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, click ⌄ in the row that contains the alarm and view the roles that are not deployed based on DR requirements in **Additional Information**.

Redeploy the role instance.

Step 3 Choose **Cluster** > **Services** > *Name of the desired service* > **Instance**. On the instance page, redeploy or adjust the role instance.

Step 4 Check whether the alarm is cleared 10 minutes later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, contact O&M personnel.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.66 ALM-12103 Executor Resource Exception

Alarm Description

HA checks the Executor resources of Manager every 30 seconds. This alarm is generated when HA detects that the Executor resources are abnormal for two consecutive times.

This alarm is cleared when the Executor resources are normal.

Resource Type of Executor is **Single-active**. Active/standby will be triggered upon resource exceptions. When this alarm is generated, the active/standby switchover is complete and new Executor resources have been enabled on the new active Manager. In this case, this alarm is cleared. This alarm is used to notify users of the cause of the active/standby Manager switchover.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12103	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- The active/standby Manager switchover occurs.
- The Executor process keeps restarting. As a result, the cluster page may fail to be accessed.

Possible Causes

The Executor process is abnormal.

Handling Procedure

Check whether the Executor process is abnormal.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command to switch to user **omm**.

Step 4 Run the **sh ${BIGDATA_HOME}/om-server/OMS/workspace0/ha/module/hacom/script/status_ha.sh** command to check whether the status of the Executor resources managed by the HA is normal. In the single-node system, the Executor resource is in the normal state. In the dual-node system, the Executor resource is in the normal state on the active node and in the stopped state on the standby node.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

Step 5 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/executor.log** command to check whether the Executor resource log of HA contains the keyword **ERROR**. If yes, analyze the log to locate the resource exception cause and fix the exception.

Step 6 After 5 minutes, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 In the **Services** area, select **Controller** and **OmmServer**, and click **OK**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.67 ALM-12104 Abnormal Knox Resources

Alarm Description

HA checks the Knox resources of Manager every 70 seconds. This alarm is generated when HA detects that the Knox resources are abnormal for three consecutive times.

This alarm is cleared when HA detects that the Knox resources are normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12104	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Requests sent by upper-layer services by using Knox cannot be properly processed.

Possible Causes

The Knox process is abnormal.

Handling Procedure

Check whether the Knox process is normal.

Step 1 Log in to FusionInsight Manager. In the alarm list, locate the row that contains the alarm and view the name of the host for which the alarm is generated.

Step 2 Use PuTTY to log in to the host for which the alarm is generated as user **root**.

Step 3 Run the **su - omm** command to switch to user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Run the **sh ${BIGDATA_HOME}/om-server/om/sbin/status-oms.sh** command to check whether the status of the Knox resources managed by HA is normal. If the status is normal, the Knox resources are normal. Otherwise, the Knox resources are abnormal.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

Step 5 Run the **vi $BIGDATA_LOG_HOME/omm/oms/ha/scriptlog/knox.log** command to check whether the Knox resource log of HA contains the keyword **ERROR**. If yes, analyze the log to locate the resource exception cause and fix the exception.

Step 6 After 5 minutes, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 In the **Services** area, select **Controller** and **OmmServer**, and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

9.11.68 ALM-12110 Failed to get ECS temporary AK/SK

Alarm Description

The meta component calls the ECS API to obtain AK/SK information every 5 minutes and caches the information. Before the AK/SK expires, the component calls the API again to update it. This alarm is generated when the component fails to call the API for three consecutive times.

This alarm is cleared when Meta successfully calls the ECS API.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12110	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The cluster cannot obtain the latest temporary AK/SK. For a storage-compute decoupled system, OBS files may fail to be accessed. As a result, upper-layer component services cannot process data.

Possible Causes

- The meta role of the MRS cluster is abnormal.
- The cluster has been bound to an agency and accessed OBS, but later it was unbound from the agency.

Handling Procedure

Check the status of the meta role.

Step 1 On FusionInsight Manager of the cluster, choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click ⌄ in the row containing the alarm, and determine the IP address of the host for which the alarm is generated.

Step 2 Choose **Cluster** > **Services** > **meta**. On the page that is displayed, click the **Instances** tab, and check whether the meta role corresponding to the host for which the alarm is generated is normal.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 Select the abnormal role, click **More**, and select **Restart Instance** to restart the abnormal meta role.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

Step 4 After the role is restarted, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Log in to the required host as user **root** using the IP address obtained in **Step 1** and check whether the **/var/log/Bigdata/meta/mrs-meta.log** file contains error information. If yes, rectify the fault based on the log information.

```
cat /var/log/Bigdata/meta/mrs-meta.log
```

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Rebind an IAM agency to the cluster.

Step 7 Log in to the MRS management console.

Step 8 In the navigation pane on the left, choose **Active Clusters**. On the page that is displayed, click the cluster name to go to its overview page. Then, check whether the cluster is bound to an IAM agency in the O&M management area.

- If yes, go to **Step 10**.
- If no, go to **Step 9**.

Step 9 Click **Select Agency**. On the page that is displayed, rebind an IAM agency that the permissions to access OBS cluster to the cluster. Then check whether the alarm is cleared a few minutes later.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager of the active cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, select **meta** for the target cluster, and click **OK**.

Step 12 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 13 Contact O&M personnel and provide the collected logs.

----End

9.11.69 ALM-12172 Failed to Report Metrics to Cloud Eye

Alarm Description

After metric sharing is enabled for a cluster, the Controller periodically collects cluster metrics and reports them to Cloud Eye.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12172	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

MRS monitoring metrics are unavailable on Cloud Eye.

Possible Causes

- Failed to call Cloud Eye APIs due to insufficient permissions.
- Failed to report data to Cloud Eye due to network problems.
- Failed to report data to Cloud Eye due to internal errors.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the additional information in the alarm details.

Step 2 Rectify the fault based on the following scenarios:

- If "Call CES to send metrics fail. Permission exception" is displayed in the additional information, the token of the resource tenant is invalid. Restart the Controller and obtain the token again. The detailed procedure is as follows:
  a. Log in to FusionInsight Manager, click the task center icon in the upper right corner, and verify that no task is being executed in the task center.
  b. Use PuTTY to log in to the active management node as user **omm**.
  c. Restart the controller.
     sh ${BIGDATA_HOME}/om-server/om/sbin/restart-controller.sh
  d. Wait for 5 minutes and check whether the alarm is cleared.
     - If yes, no further action is required.
     - If no, go to **Step 3**.
- If "Call CES to send metrics fail. Request CES error code xxx" or "CES internal error code xxx" is displayed, the function service is abnormal. Go to **Step 3**.
- If "Call CES to send metrics fail. Too many request" is displayed in additional alarm information, the service request triggers triffic control. Go to **Step 3**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 4 On the displayed page, select **Controller** of **OMS** for **Service**, and click **OK**. Select the active node for **Host** and click **OK**.

Step 5 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 6 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.70 ALM-12180 Suspended Disk I/O

Alarm Description

For MRS 3.3.0 and its later versions:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 6 seconds within 30 seconds in at least seven collection periods.
  - The system collects data every 3 seconds by default, and within 30 seconds, at least 10 collection cycles have a disk queue depth (**avgqu-sz**) > 0, IOPS = 0, or bandwidth = 0, and **ioutil** > 99%.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 1000 ms within 300 seconds.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 3 seconds within 30 seconds in at least seven collection periods.
  - By default, the system collects data every 3 seconds. The disk queue depth (**avgqu-sz**) is greater than 0, the IOPS or bandwidth is 0, and **ioutil** is greater than 99% in at least 10 collection periods within 30 seconds.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 500 ms within 300 seconds.

The collection period is 3 seconds, and the detection period is 30 or 300 seconds. This alarm is automatically cleared when neither of the preceding conditions is met for three consecutive detection periods (30 or 300 seconds).

For versions earlier than MRS 3.3.0:

- For HDDs, the alarm is triggered when any of the following conditions is met:

  - By default, the system collects data every 3 seconds. The svctm latency exceeds 6 seconds within 30 seconds in at least 10 collection periods.
  - By default, the system collects data every 3 seconds. The disk queue depth (**avgqu-sz**) is greater than 0, the IOPS or bandwidth is 0, and **ioutil** is greater than 99% in at least 10 collection periods within 30 seconds.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency exceeds 2 seconds within 30 seconds in at least 10 collection periods.
  - By default, the system collects data every 3 seconds. The disk queue depth (**avgqu-sz**) is greater than 0, the IOPS or bandwidth is 0, and **ioutil** is greater than 99% in at least 10 collection periods within 30 seconds.

This alarm is automatically cleared when the preceding conditions have not been met for 90s.

NOTE

For details about how to obtain and calculate related parameters, see **Related Information**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12180	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
DiskName	Specifies the disk for which the alarm was generated.

Impact on the System

If the I/O usage keeps increasing, operations will be affected and services will be interrupted. The possible impacts are as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The system I/O performance deteriorates, which means slow response and low throughput. For example, job submission is slow, page responds slowly, interface response times out, and the system is in error or even crash.
- Customer services may be interrupted. The system may break down and the key information stored on the faulty disk may be lost.

Possible Causes

The disk is aged.

Handling Procedure

Replace the disk.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**.

Step 2 View the detailed information about the alarm. Check the values of **HostName** and **DiskName** in the location information to obtain the information about the faulty disk for which the alarm is reported.

Step 3 Replace the hard disk.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Select **OMS** for **Service** and click **OK**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

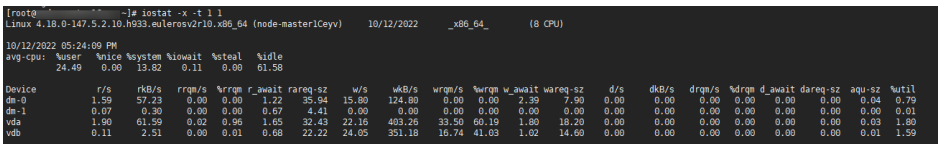
This alarm is automatically cleared after the fault is rectified.

Related Information

Obtain and calculate related parameters as follows:

- Run the following command in the OS to collect data:

  iostat -x -t 1 1



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters are as follows:

avgqu-sz indicates the disk queue depth.

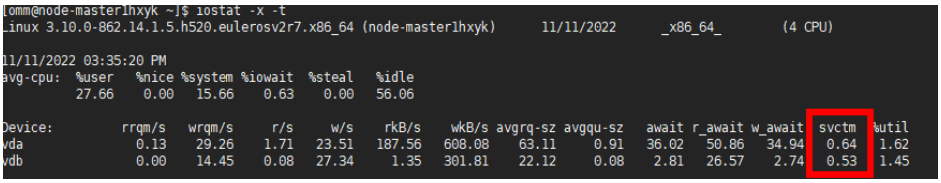
The sum of **r/s** and **w/s** is the IOPS.

The sum of **rkB/s** and **wkB/s** is the bandwidth.

%util is the **ioutil** value.

- MRS 3.1.0:

  Run the **iostat -x -t** command in the OS.



- Calculate **svctm** as follows in versions later than MRS 3.1.0:

  svctm = (tot_ticks_new - tot_ticks_old) / (rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old)

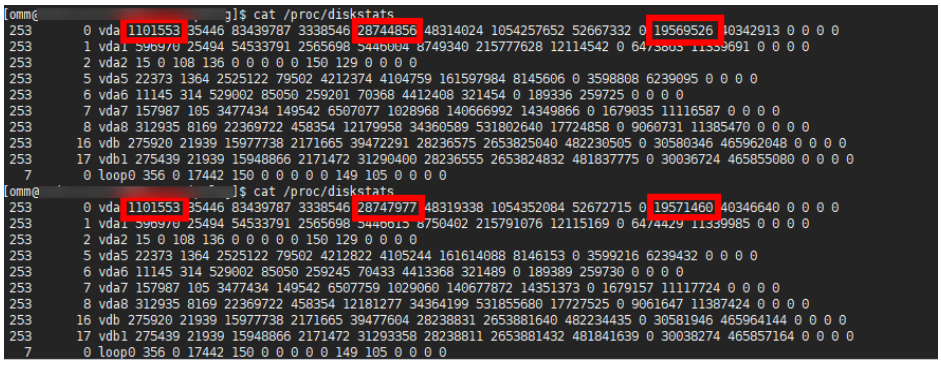
- Versions earlier than MRS 3.3.0: If **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, then **svctm = 0**.
- MRS 3.3.0 and its later versions:

  When the detection period is 30 seconds, if **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, then **svctm = 0**.

  When the detection period is 300 seconds and **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, if **tot_ticks_new - tot_ticks_old = 0**, then **svctm = 0**; otherwise, the value of **svctm** is infinite.

The parameters can be obtained as follows:

The system runs the **cat /proc/diskstats** command every 3 seconds to collect data. For example:



In the two commands:

In the data collected for the first time, the number in the fourth column is the **rd_ios_old** value, the number in the eighth column is the **wr_ios_old** value, and the number in the thirteenth column is the **tot_ticks_old** value.

In the data collected for the second time, the number in the fourth column is the **rd_ios_new** value, the number in the eighth column is the **wr_ios_new** value, and the number in the thirteenth column is the **tot_ticks_new** value.

In this case, the value of **svctm** is as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

(19571460 - 19569526) / (1101553 + 28747977 - 1101553 - 28744856) = 0.6197

9.11.71 ALM-12186 CGroup Task Usage Exceeds the Threshold

Alarm Description

The system checks the CGroup task usage of user **omm** every 5 minutes. This alarm is generated when the CGroup task usage exceeds 90%. This alarm is cleared when the CGroup task usage is less than or equal to 90%.

CGroup task usage = Number of used CGroup tasks/Maximum number of CGroup tasks

You can run the **systemctl status user-$(id -u).slice | grep limit | awk -F ' ' '{print $2}'** command as user **omm** to obtain the number of used CGroup tasks of this user and run the **echo $(systemctl status user-$(id -u).slice | grep limit | awk -F ' ' '{print $4}') | sed -e 's/)//g'** command to obtain the maximum number of CGroup tasks allowed for this user.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12186	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- Failed to switch to user **omm**.
- Failed to create new **omm** processes.
- A faulty service or process cannot be restarted.

Possible Causes

The CGroup task usage exceeds 90%.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the maximum number of threads that can be concurrently opened by user omm is properly set.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click ⌄ in the row containing the alarm, and view the name of the host for which the alarm is generated in **Location**. Click the host name to view its IP address.

Step 2 Log in to the host for which the alarm is generated as user **omm**.

Step 3 Run the following command to obtain the maximum number of threads that can be concurrently opened by user **omm** and check whether this number is greater than or equal to **60000**:

systemctl status user-$(id -u).slice | grep limit

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Switch to user **root** and run the following command to change the value for user **omm** to **60000**:

systemctl set-property user-2000.slice TasksMax=60000

Step 5 Change the value of **UserTasksMax** in the **/etc/systemd/logind.conf** file to **60000**. (If the parameter is commented out, uncomment it.) Save the file, wait 5 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager of the cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, select **OmmServer** and **NodeAgent** for the target cluster, and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.72 ALM-12187 Failed to Expand Disk Partition Capacity

Alarm Description

The system checks the disk space every 60 seconds. When detecting that the disk space is expanded, the system expands disk partition. This alarm is generated when the disk partition fails to be expanded.

This alarm is cleared when the system detects that the disk partition is successfully expanded.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12187	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
MountDirectoryName	Specifies the directory for which the alarm is generated.

Impact on the System

A disk partition scale-out failure may have the following impacts on the system:

- Data may be lost. (It is important to back up key information before scale-out.)
- The system may be unstable or cannot be started because the system files may be damaged.
- The disk may be unavailable. In this case, you need to format the disk and partition it again.
- The system performance may deteriorate because the disk partition space is not enough. The performance remains low until the disk partition is scaled out successfully.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The growpart scale-out tool is not installed.
- The system fails to execute the command for expanding disk partition.

Handling Procedure

Check whether growpart is installed.

Step 1 Log in to FusionInsight Manager, click **O&M**, and choose **Alarm** > **Alarms** to view the alarm details. In the **Location** column, check the name of the host and mount directory for which the alarm is generated. Click the host name to view its IP address.

Step 2 Log in to the node for which the alarm is generated as user **root**.

Step 3 Run the following command to check whether growpart is installed:

which growpart

If information similar to the following is displayed, the growpart tool is installed. Otherwise, contact O&M personnel to install the growpart tool.

```
[root@xxx ~]#which growpart
/usr/bin/growpart
```

Step 4 Wait for 5 minutes, then choose **O&M**, and choose **Alarm** > **Alarms** on FusionInsight Manager. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Run the disk partition expansion command.

Step 5 Run the following command to view the disk and partition information:

lsblk

Search for the partition and the disk based on the mount directory name in the alarm location information, and check the disk and partition sizes.

In the following example, the mount directory is **/srv/BigData/data1**, the used disk is **/dev/vdb**, and the disk partition is **/dev/vdb1**.

```
[root@                    ~]# lsblk
NAME   MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
loop0    7:0    0  1.5G  0 loop /media
vda    253:0    0  500G  0 disk
├─vda1 253:1    0  220G  0 part /
├─vda2 253:2    0    1K  0 part
├─vda5 253:5    0   10G  0 part /tmp
├─vda6 253:6    0   10G  0 part /var
├─vda7 253:7    0   60G  0 part /srv/BigData
└─vda8 253:8    0  180G  0 part /var/log
vdb    253:16   0  650G  0 disk
└─vdb1 253:17   0  600G  0 part /srv/BigData/data1
```

Step 6 Run the following command to expand the partition using growpart:

growpart *Data disk Partition number*

Run the following command:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

growpart /dev/vdb 1

If information similar to the following is displayed, the execution is successful. If the execution fails, contact O&M personnel.

```
[root@host-                    ~]# growpart /dev/vdb 1
CHANGED: partition=1 start=2048 old: size=1258287104 end=1258289152 new: size=1363146719 end=1363148767
```

Step 7 Run the following command to expand the file system size of the disk partition:

resize2fs *Disk partition*

Run the following command:

resize2fs /dev/vdb1

If information similar to the following is displayed, the execution is successful:

```
[root@host-                    ~]# resize2fs /dev/vdb1
resize2fs 1.46.4 (18-Aug-2021)
Filesystem at /dev/vdb1 is mounted on /srv/BigData/data1; on-line resizing required
old_desc_blocks = 75, new_desc_blocks = 82
The filesystem on /dev/vdb1 is now 170393339 (4k) blocks long.
```

Step 8 Wait for 5 minutes, click **O&M**, and choose **Alarm** > **Alarms** on FusionInsight Manager. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, contact O&M personnel.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.73 ALM-12188 diskmgt Disk Monitoring Unavailable

Alarm Description

NodeAgent checks the status of the diskmgt disk monitoring service every 5 minutes. This alarm is generated when diskmgt disk monitoring is unavailable.

This alarm is cleared when the diskmgt disk monitoring service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12188	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

When diskmgt disk monitoring is unavailable, the read-only detection of the device partition file system, device partition loss detection, and disk partition scale-out detection cannot be performed.

Possible Causes

- The diskmgt disk monitoring service does not exist.
- The diskmgt disk monitoring service is not started.

Handling Procedure

Check whether the diskmgt disk monitoring service exists.

Step 1 Log in to FusionInsight Manager, click **O&M**, and choose **Alarm** > **Alarms** to view the alarm details. In the **Location** column, check the name of the host for which the alarm is generated. Click the host name to view its IP address.

Step 2 Log in to the node for which the alarm is generated as user **root**.

Step 3 Run the following command to check whether the core service file exists:

stat /usr/local/diskmgt/inner/diskmgtd

If the file does not exist, contact O&M personnel.

Start the diskmgt disk monitoring service.

Step 4 Run the following command to start the diskmgt disk monitoring service:

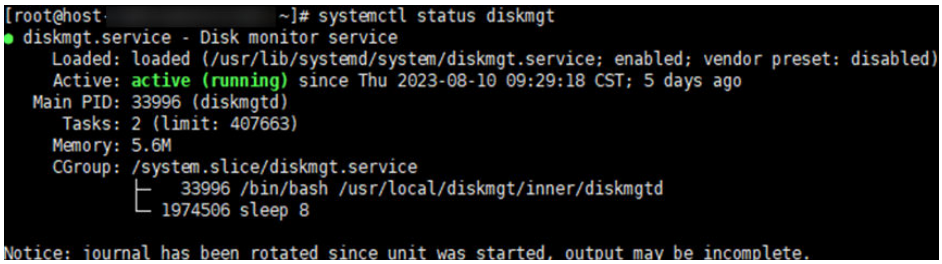
systemctl restart diskmgt

Step 5 Run the following command to check whether the diskmgt disk monitoring service is started:

systemctl status diskmgt

- If information similar to the following is displayed, the service is started successfully. Go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



- If no, contact O&M personnel.

Step 6 Wait for 5 minutes, click **O&M**, and choose **Alarm** > **Alarms** on FusionInsight Manager. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, contact O&M personnel.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.74 ALM-12190 Number of Knox Connections Exceeds the Threshold

Alarm Description

The system periodically checks the number of connections to all Knox topologies. This alarm is generated when the number of connections to a topology exceeds the threshold (90% by default). This alarm is automatically cleared when the number of connections to a topology falls below the threshold.

NOTE

This alarm applies to clusters of MRS 3.1.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12190	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Topology	Specifies the Knox topology for which the alarm was generated.

Impact on the System

The topology may reach the upper limit of connections and fail to forward requests, adversely affecting the MRS functions.

Possible Causes

Hue or Manager is too frequently used, but the default maximum number of Knox connections is small.

Handling Procedure

Step 1 Log in to active and standby OMS nodes as user **root**, respectively.

Step 2 Add the following configuration to the **gateway-site.xml** file on the active and standby OMS nodes to increase the number of thread pools:

vi /opt/knox/conf/gateway-site.xml

```
<property>
<name>gateway.httpclient.maxConnections</name>
<value>64</value>
</property>
```

Step 3 Log in to the active OMS node as user **omm** and run the following command to restart the Knox process:

sh /opt/knox/bin/restart-knox.sh

Step 4 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Contact O&M personnel to rectify the fault.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.75 ALM-12191 Disk I/O Usage Exceeds the Threshold

Alarm Description

The system checks the disk I/O usage every 30 seconds and compares the actual disk I/O usage with the threshold. This alarm is generated when the disk I/O usage exceeds the threshold for multiple consecutive times (**10** by default).

If the **hit number** is **1**, this alarm is cleared when the disk I/O usage is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the disk I/O usage is less than or equal to 90% of the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12191	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- Latency: Service processes may run slowly and there is a latency.
- Service failure: Service processing may be slow, time out, or fail. As a result, jobs may fail to run.

Possible Causes

- The alarm threshold or alarm trigger count is improperly configured.
- The disk configuration cannot meet service requirements. The disk I/O usage reaches the upper limit. Alternatively, services are in peak hours. The disk I/O usage reaches the upper limit in a short period.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Modify the alarm threshold and alarm trigger count based on the actual disk I/O usage.

1. Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Host** > **Disk** > **Disk IO Utilization**.
2. Click the edit button next to **Trigger Count** to change it to a proper value based on the actual service usage.

   NOTE

   Trigger Count indicates how many consecutive times the threshold is reached when the alarm is triggered.
3. Click **Modify** in the **Operation** column of the row that contains the rule and change the alarm threshold.

Step 2 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check whether the disk I/O usage reaches the upper limit.

Step 3 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details and click the name of the host for which the alarm is generated in **Location** area.

Step 4 On the overview page of the host, observe the real-time data of the disk I/O usage for about 5 minutes. If the disk I/O usage exceeds the threshold for multiple times, contact the MRS cluster administrator to improve the disk specification.

If **Disk IO Utilization** chart is not displayed, click the drop-down arrow on the right, select **Customize**, select the desired item, and click **OK**.

Step 5 Check whether it was the peak hour. If this alarm was generated during peak hours, expand the node capacity or contact the MRS cluster administrator to improve the disk specification.

Step 6 Check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, select **NodeAgent** for the target cluster, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.76 ALM-12192 Host Load Exceeds the Threshold

Alarm Description

The system checks the average load every 30 seconds and compares the actual average load with the threshold. This alarm is generated when the average load exceeds the threshold for multiple consecutive times (10 by default).

This alarm is cleared when **Trigger Count** is **1** and the average load is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the average load is less than or equal to 90% of the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12192	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

- Latency: Service processes may run slowly and there is a latency.
- Service failure: Service processing may be slow, time out, or fail. As a result, jobs may fail to run.

Possible Causes

The host cannot meet service requirements. The average load reaches the upper limit. Alternatively, requirements surged during peak hours, and the average load reaches the upper limit in a short period.

Handling Procedure

Check the host CPU load.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details and click the name of the host for which the alarm is generated in **Location** area.

Step 2 On the **Hosts** page, select the host for which the alarm is generated. Click the **Chart** tab, select **Host Status**, and check whether the **Average Host Load per CPU Core** is greater than 3.

- If yes, the system is overloaded. Go to **Step 3**.
- If no, go to **Step 8**.

Step 3 Log in to the host for which the alarm is reported as user **omm**.

Step 4 Run the **top** command to check whether the **us** value of **%Cpu(s)** is greater than 80.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
top - 16:12:33 up 8 days,  4:51,  9 users,  load average: 18.87, 19.97, 20.89
Tasks: 1190 total,   7 running, 1183 sleeping,   0 stopped,   0 zombie
%Cpu(s): 38.5 us, 37.3 sy,  0.0 ni, 24.0 id,  0.0 wa,  0.0 hi,  0.3 si,  0.0 st
KiB Mem : 26357323+total, 19515480 free, 10638136+used, 13767638+buff/cache
KiB Swap:        0 total,        0 free,        0 used. 14276078+avail Mem

  PID USER      PR  NI    VIRT    RES    SHR S  %CPU %MEM     TIME+ COMMAND
20465 omm       20   0 6474204 145540  23680 S 212.6  0.1   0:06.59 java
 2081 omm       20   0   21.2g   2.4g  18596 S 104.5  0.9   3870:30 java
 5909 omm       20   0 5983500 824128 105800 S 100.3  0.3 136:47.44 java
18519 omm       20   0   17.1g   3.2g  28640 S  99.0  1.3  10313:24 java
32531 omm       20   0   36.9g 130096  53004 S  51.0  0.0   0:01.58 java
```

- If yes, the CPU usage of user processes is too high. Record the *PIDs* of the processes with high CPU usage and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Run the following command to obtain the name of the process with high CPU usage, query the process logs (in the **/var/log/Bigdata** directory) based on the process name, and check whether the service logs contain error information:

ps -ef | grep *PID*

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 Run the **top** command to check whether the value of **wa** is greater than 10.0.

```
top - 16:06:37 up 7 days, 23:26,  7 users,  load average: 8.58, 9.00, 8.18
Tasks: 486 total,   1 running, 485 sleeping,   0 stopped,   0 zombie
%Cpu(s): 56.0 us, 11.2 sy,  0.0 ni, 32.2 id,  0.0 wa,  0.4 hi,  0.3 si,  0.0 st
MiB Mem :  64866.6 total,   7948.8 free,  35870.1 used,  21047.7 buff/cache
MiB Swap:      0.0 total,      0.0 free,      0.0 used.  24219.1 avail Mem

   PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM     TIME+ COMMAND
304130 omm       20   0   12.7g   1.9g 100032 S 125.5   3.0  50:55.11 java
2388328 omm      20   0   21.0g 110464  35776 S  84.8   0.2   0:02.56 java
2387356 omm      20   0 1096000 169152  53184 S  76.8   0.3   0:02.32 java
```

- If yes, the read and write performance reaches the bottleneck. Go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Run the **iotop** command as the **root** user to check the processes with high disk read/write usage and determine whether the processes are unnecessary based on service needs.

- If yes, run the following command to stop unnecessary processes. (If PID is not displayed, press **P** to switch TID to PID.)

  kill -9 *PID*

- If no, go to **Step 8**.

Step 8 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Expand the **Service** drop-down list, select **NodeAgent** for the target cluster, and click **OK**.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.77 ALM-12200 Password Is About to Expire

Alarm Description

The system checks whether a user password is about to expire at 1:00 a.m. every day. This alarm is generated when a user password is about to expire in less than 5 days by default.

This alarm is cleared when the password is about to expire in at least five days by default.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12200	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
Additional Information	Details	Specifies that the username of password that is about to expire.

Impact on the System

The account cannot be used.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The password is about to expire.

Handling Procedure

Change the user password.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details, and view and record the name of the user whose password is about to expire in additional information.

Step 2 Change the password.

Step 3 If DataArts Studio is interconnected, check whether DataArts Studio jobs are using a password that is about to expire. If yes, go to the DataArts Studio management center to change the password. Otherwise, a large number of jobs may fail.

Step 4 Check whether the alarm is automatically cleared after 1:00 a.m. the next day.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Select **Controller** for **Service** and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.78 ALM-12201 Process CPU Usage Exceeds the Threshold

Alarm Description

The system checks the CPU usage every 30 seconds and compares the check result with the default threshold. This alarm is generated when the CPU usage exceeds the threshold for multiple consecutive times (**10** by default).

This alarm is cleared when **Trigger Count** is **1** and the CPU usage is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the CPU usage is less than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12201	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

- Latency: Service processes may run slowly and there is a latency.
- Service failure: Service processing may be slow, time out, or fail. As a result, jobs may fail to run.

Possible Causes

- The alarm threshold or alarm trigger count is improperly configured.
- The CPU configuration cannot meet service requirements, and the CPU usage reaches the upper limit. Alternatively, services are in peak hours. The CPU usage reaches the upper limit in a short period.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Modify the alarm threshold and alarm trigger count based on the actual CPU usage.

1. Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds** > **OMS** > **OMSServices** > **CPU** > **Process Used CPU (OMS)**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Click the edit button next to **Trigger Count** to set it a proper value based on the actual service usage.

   NOTE

   Trigger Count indicates how many consecutive times the threshold is reached when the alarm is triggered.

3. Click **Modify** in the **Operation** column of the row that contains the rule and change the alarm threshold.

Step 2 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check whether the CPU usage reaches the upper limit.

Step 3 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details and click the name of the host for which the alarm is generated in **Location** area.

Step 4 On the overview page of the host, observe the real-time data of the host CPU usage for about 5 minutes. If the CPU usage exceeds the threshold for multiple times, contact the MRS cluster administrator to increase the CPU.

If no chart is available, click the drop-down arrow on the right, select **Customize**, select the desired item, and click **OK**.

Step 5 Check whether it was the peak hour. If this alarm was generated during peak hours, expand the node capacity or contact the MRS cluster administrator to improve the CPU specification.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, select **OmmServer** for the target cluster, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None.

9.11.79 ALM-12202 Process Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of main OMS processes every 30 seconds. This alarm is generated when the memory usage of main OMS processes is greater than 90% (default value) of the maximum memory.

This alarm is cleared when the memory usage of main OMS processes is less than or equal to 90% of the maximum memory.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12202	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

If the memory usage of main OMS processes is too high, the performance of these processes deteriorates, and even memory overflow occurs. As a result, main OMS processes are unavailable, and OMS tasks are slow or fail to run.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The memory usage of main OMS processes is too high or the memory is inappropriately allocated.

Handling Procedure

Check the memory usage of main OMS processes.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details, record the process name in **Location**, click the reported host name, and record the service IP address of the host.

Step 2 Choose **System** > **OMS** to view the **OMS Process Memory Usage Ratio** chart. Check whether the memory usage of the processes reaches the threshold (90% by default) at the time when the alarm is generated.

If no chart is available, click the drop-down arrow on the right, select **Customize**, select the desired item, and click **OK**.

- If yes, go to **Step 3**.
- If the threshold is not reached, go to **Step 6**.

Step 3 Contact O&M personnel to modify the memory configurations of the processes.

Step 4 Restart the processes for which the alarm is generated.

Step 5 Check whether the alarm is cleared in 10 minutes.

- If yes, no further action is required.
- If the threshold is not reached, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **OmmServer** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.80 ALM-12203 Process Full GC Duration Exceeds the Threshold

Alarm Description

The system checks the GC duration of main OMS processes every 30 seconds. If the GC duration of an OMS process exceeds the threshold for three consecutive times, this alarm is generated. You can choose **O&M** > **Alarm** > **Thresholds** > **OMS** > **OMSServices** to change the threshold.

This alarm is cleared when the GC duration of the OMS process is shorter than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12203	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger condition	Specifies the alarm triggering condition.

Impact on the System

Read and write performance deteriorates. As a result, the task execution may slow down and even the service may restart unexpectedly.

Possible Causes

The memory of main OMS processes is too high or inappropriately allocated, causing frequent occurrence of the full GC.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details, record the process name in **Location**, click the reported host name, and record the service IP address of the host.

Step 2 Choose **System** > **OMS**, view the **Full GC Duration of an OMS Process** - *Process Name* chart, and check whether the GC duration is longer than 12 seconds (default value).

If no chart is available, click the drop-down arrow on the right, select **Customize**, select the desired item, and click **OK**.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Contact O&M personnel to modify the memory configurations of the processes.

Step 4 Restart the process.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **OmmServer** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.81 ALM-12204 Wait Duration of a Disk Read Exceeds the Threshold

Alarm Description

The system checks the wait duration of a disk read every 30 seconds and compares the actual wait duration with the threshold. This alarm is generated

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

when the wait duration exceeds the threshold (10s by default) for multiple consecutive times.

This alarm is cleared when the wait duration is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12204	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

- Latency: Service processes may run slowly and there is a latency.
- Service failure: Service processing may be slow, time out, or fail. As a result, jobs may fail to run.

Possible Causes

- The alarm threshold or alarm trigger count is improperly configured.
- The disk configuration cannot meet service requirements. The disk I/O performance reaches the upper limit. Alternatively, services are in peak hours. The wait duration of a disk read reaches the upper limit in a short period.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Modify the alarm threshold and alarm trigger count based on the actual disk I/O usage.

1. Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Disk** > **Average Time Required for Each Read operation**.
2. Click the edit button next to **Trigger Count** to set it a proper value based on the actual service usage.

   NOTE

   Trigger Count indicates how many consecutive times the threshold is reached when the alarm is triggered.
3. Click **Modify** in the **Operation** column of the row that contains the rule and change the alarm threshold.

Step 2 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check whether the average time required for each read operation reaches the upper limit.

Step 3 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details and click the name of the host for which the alarm is generated in **Location** area.

Step 4 On the overview page of the host, observe the real-time data of average time required for each read operation for about 5 minutes. If the wait duration exceeds the threshold for multiple times, contact the MRS cluster administrator to improve the disk specification.

If the **Average Time Required for Each Read Operation** chart is unavailable, click the drop-down arrow on the right, select **Customize**, select the corresponding item, and click **OK**.

Step 5 Check whether it was the peak hour. If this alarm was generated during peak hours, expand the node capacity or contact the MRS cluster administrator to improve the disk specification.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, select **NodeAgent** for the target cluster, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.82 ALM-12205 Wait Duration of a Disk Write Exceeds the Threshold

Alarm Description

The system checks the wait duration of a disk write every 30 seconds and compares the actual wait duration with the threshold. This alarm is generated when the wait duration exceeds the threshold (10s by default) for multiple consecutive times.

This alarm is cleared when the wait duration is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12205	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- Latency: Service processes may run slowly and there is a latency.
- Service failure: Service processing may be slow, time out, or fail. As a result, jobs may fail to run.

Possible Causes

- The alarm threshold or alarm trigger count is improperly configured.
- The disk configuration cannot meet service requirements. The disk I/O performance reaches the upper limit. Alternatively, services are in peak hours. The wait duration of a disk write reaches the upper limit in a short period.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Modify the alarm threshold and alarm trigger count based on the actual disk I/O usage.

1. Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Host** > **Disk** > **Average Time Required for Each Write Operation**.
2. Click the edit button next to **Trigger Count** to set it a proper value based on the actual service usage.

   NOTE

   Trigger Count indicates how many consecutive times the threshold is reached when the alarm is triggered.
3. Click **Modify** in the **Operation** column of the row that contains the rule and change the alarm threshold.

Step 2 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check whether the average time required for each write operation reaches the upper limit.

Step 3 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details and click the name of the host for which the alarm is generated in **Location** area.

Step 4 On the overview page of the host, observe the real-time data of average time required for each write operation for about 5 minutes. If the wait duration exceeds the threshold for multiple times, contact the MRS cluster administrator to improve the disk specification.

If the **Average Time Required for Each Write Operation** chart is not displayed, click the drop-down arrow on the right, select **Customize**, select the desired item, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Check whether it was the peak hour. If this alarm was generated during peak hours, expand the node capacity or contact the MRS cluster administrator to improve the disk specification.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, select **NodeAgent** for the target cluster, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.83 ALM-12206 Password Has Expired

Alarm Description

The system checks whether a user password has expired at 1:00 a.m. every day. This alarm is generated when a user password has expired.

This alarm is cleared when the user password in the system is within the validity period.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12206	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
Additional Information	Details	Specifies that the username of password that has expired.

Impact on the System

The account cannot be used.

Possible Causes

The user password has expired.

Handling Procedure

Change the user password.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details, and view and record the name of the user whose password has expired in additional information.

Step 2 Change the user password that has expired.

Step 3 If DataArts Studio is interconnected, check whether DataArts Studio jobs are using an expired user password. If yes, go to the DataArts Studio management center to change the password and execute the affected jobs again.

Step 4 Check whether the alarm is automatically cleared after 1:00 a.m. the next day.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Select **Controller** for **Service** and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.84 ALM-12207 Slow Disk Processing Timeout

Alarm Description

When slow disk detection is enabled, the system checks the slow disk processing status every 10 minutes by default. This alarm is generated when the following disk or node status does not change within 10 hours.

Disk: Automatic isolation aborted, isolated, isolation failed, and de-isolation failed.

Node: Isolated, Isolation failed, Isolation cancellation failed, Node startup failed, and De-isolated.

This alarm is automatically cleared when the status of the node or disk that is in the processing timeout state changes.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
12207	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
	DiskName	Specifies the disk for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
Additional Information	HostName	Specifies the host for which the alarm was generated.
	DiskName	Specifies the disk for which the alarm was generated.
	Details	Specifies that the description of slow disk isolation.

Impact on the System

If an isolated disk or node cannot be restored in a timely manner, the running of components may be affected, which further affects user services.

Possible Causes

The isolation status of the disk or node exceeds the configured timeout period for processing slow disks.

Handling Procedure

Check the cause of the slow disk processing timeout.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the alarm list, expand the alarm details, and view and record the host or disk for which the alarm is generated.

Step 2 Log in to the active OMS node as user **root** and run the following command to check the cause of slow disk processing timeout in the controller log and check whether there is obvious error information:

vi /var/log/Bigdata/controller/controller.log

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 Log in to the node for which the alarm is generated as user **root** and run the following command to check the cause of slow disk processing timeout in the agent log and check whether any error information is displayed:

vi /var/log/Bigdata/nodeagent/agentlog/agent.log

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Contact O&M personnel to rectify the fault and manually run the command for the slow disk or node. After the command is executed, observe for 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Select **Controller** and **NodeAgent** for **Service**, select the active/standby OMS node and the node for which the alarm is generated in the **Host** area, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.85 ALM-13000 ZooKeeper Service Unavailable

Description

The system checks the ZooKeeper service status every 60 seconds. This alarm is generated when the ZooKeeper service is unavailable.

This alarm is cleared when the ZooKeeper service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13000	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

ZooKeeper cannot provide coordination services for upper layer components and the components (such as Yarn and Flink) that depend on ZooKeeper may not run properly.

Possible Causes

- The DNS is installed on the ZooKeeper node.
- The network is faulty.
- The KrbServer service is abnormal.
- The ZooKeeper instance is abnormal.
- The disk capacity is insufficient.

Procedure

Check the DNS.

Step 1 Check whether the DNS is installed on the node where the ZooKeeper instance is located. On the Linux node where the ZooKeeper instance is located, run the **cat /etc/resolv.conf** command to check whether the file is empty.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Run the **service named status** command to check whether the DNS is started.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Run the **service named stop** command to stop the DNS service. If "Shutting down name server BIND waiting for named to shut down (28s)" is displayed, the DNS service is stopped successfully. Comment out the content (if any) in **/etc/resolv.conf**.

Step 4 On the **O&M > Alarm > Alarms** tab, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the network status.

Step 5 On the Linux node where the ZooKeeper instance is located, run the **ping** command to check whether the host names of other nodes where the ZooKeeper instance is located can be pinged successfully.

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 Modify the IP addresses in **/etc/hosts** and add the host name and IP address mapping.

Step 7 Run the **ping** command again to check whether the host names of other nodes where the ZooKeeper instance is located can be pinged successfully.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 8**.
- If no, go to **Step 23**.

Step 8 On the **O&M > Alarm > Alarms** tab, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check the KrbServer service status (Skip this step if the normal mode is used).

Step 9 On FusionInsight Manager, choose **Cluster >** *Name of the desired cluster* **> Services**.

Step 10 Check whether the KrbServer service is normal.

- If yes, go to **Step 13**.
- If no, go to **Step 11**.

Step 11 Rectify the fault by referring to **ALM-25500 KrbServer Service Unavailable** and check whether the KrbServer service is restored.

- If yes, go to **Step 12**.
- If no, go to **Step 23**.

Step 12 On the **O&M > Alarm > Alarms** tab, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Check the ZooKeeper service instance status.

Step 13 On FusionInsight Manager, choose **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper** > **quorumpeer**.

Step 14 Check whether the ZooKeeper instances are normal.

- If yes, go to **Step 18**.
- If no, go to **Step 15**.

Step 15 Select instances whose status is not good, and choose **More** > **Restart Instance**.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform this operation during off-peak hours.

Step 16 Check whether the instance status is good after restart.

- If yes, go to **Step 17**.
- If no, go to **Step 18**.

Step 17 On the **O&M > Alarm > Alarms** tab, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Check disk status.

Step 18 On FusionInsight Manager, choose **Cluster >** *Name of the desired cluster* **> Service** > **ZooKeeper** > **quorumpeer**, and check the node host information of the ZooKeeper instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 19 On FusionInsight Manager, click **Host**.

Step 20 In the **Disk** column, check whether the disk space of each node where ZooKeeper instances are located is insufficient (disk usage exceeds 80%).

- If yes, go to **Step 21**.
- If no, go to **Step 23**.

Step 21 Expand disk capacity. For details, see **ALM-12017 Insufficient Disk Capacity**.

Step 22 On the **O&M > Alarm > Alarms** tab, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 23**.

Collect fault information.

Step 23 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 24 Select the following nodes in the required cluster from the **Service**: (KrbServer logs do not need to be downloaded in normal mode.)

- ZooKeeper
- KrbServer

Step 25 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 26 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.86 ALM-13001 Available ZooKeeper Connections Are Insufficient

Description

The system checks ZooKeeper connections every 60 seconds. This alarm is generated when the system detects that the number of used ZooKeeper instance connections exceeds the threshold (80% of the maximum connections).

When the **Trigger Count** is 1, this alarm is cleared when the number of used ZooKeeper instance connections is smaller than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the number of used ZooKeeper instance connections is smaller than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13001	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host name for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Available ZooKeeper connections are insufficient. When the connection usage reaches 100%, the ZooKeeper cannot process external connections. As a result, upstream components (such as Yarn and Flink) cannot run properly.

Possible Causes

The number of connections to the ZooKeeper node exceeds the threshold. Connection leakage occurs on some connection processes, or the maximum number of connections does not comply with the actual scenario.

Procedure

Check connection status.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms**. On the displayed interface, click the drop-down button of **Available ZooKeeper Connections Are Insufficient** and confirm the node IP address of the host for which the alarm is generated in the Location Information.

Step 2 Obtain the PID of the ZooKeeper process. Log in to the node involved in this alarm as user **root** and run the **pgrep -f proc_zookeeper** command.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the PID can be correctly obtained.

- If yes, go to **Step 4**.
- If no, go to **Step 15**.

Step 4 Obtain all the IP addresses connected to the ZooKeeper instance and the number of connections and check 10 IP addresses with top connections. Run the following command based on the obtained PID: **lsof -i|grep** *$pid* **| awk '{print $9}' | cut -d : -f 2 | cut -d \>-f 2 | awk '{a[$1]++} END {for(i in a){print i,a[i] | "sort -r -g -k 2"}}' | head -10**. (The PID obtained in the preceding step is used.)

Step 5 Check whether node IP addresses and number of connections are successfully obtained.

- If yes, go to **Step 6**.
- If no, go to **Step 15**.

Step 6 Obtain the ID of the port connected to the process. Run the following command based on the obtained PID and IP address: **lsof -i|grep** *$pid* **| awk '{print $9}'|cut -d \> -f 2 |grep** *$IP*| **cut -d : -f 2**. (The PID and IP address obtained in the preceding step are used.)

Step 7 Check whether the port ID is successfully obtained.

- If yes, go to **Step 8**.
- If no, go to **Step 15**.

Step 8 Obtain the ID of the connected process. Log in to each IP address and run the following command based on the obtained port ID: **lsof -i|grep** *$port*. (The port ID obtained in the preceding step is used.)

Step 9 Check whether the process ID is successfully obtained.

- If yes, go to **Step 10**.
- If no, go to **Step 15**.

Step 10 Check whether connection leakage occurs on the process based on the obtained process ID.

- If yes, go to **Step 11**.
- If no, go to **Step 12**.

Step 11 Close the process where connection leakage occurs and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Step 12 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper** > **Configurations** > **All Configurations** > **quorumpeer** > **Performance** and increase the value of **maxCnxns** as required.

Step 13 Save the configuration and restart the ZooKeeper service.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 14 Check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 15**.

Collect fault information.

Step 15 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 16 Select **ZooKeeper** in the required cluster from the **Service**:

Step 17 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 18 Contact the O&M personnel and send the collected log information.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.87 ALM-13002 ZooKeeper Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the ZooKeeper service every 30 seconds. The alarm is generated when the direct memory usage of a ZooKeeper instance exceeds the threshold (80% of the maximum memory).

When the **Trigger Count** is 1, this alarm is cleared when the ZooKeeper Direct memory usage is less than the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the ZooKeeper Direct memory usage is less than 80% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13002	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available memory of ZooKeeper is insufficient, memory overflow may occur and services may break down. As a result, upstream services (such as HDFS and Yarn) fail to run.

Possible Causes

The direct memory of the ZooKeeper instance is overused or the direct memory is inappropriately allocated.

Procedure

Check the direct memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms**. On the displayed interface, click the drop-down button of **ZooKeeper Direct Memory Usage Exceeds the Threshold**. Check the IP address of the instance that reports the alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper** > **Instance** > **quorumpeer(the IP address checked)**. Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **CPU and Memory**, and select **ZooKeeper Heap And Direct Buffer Resource Percentage**, click **OK**.

Step 3 Check whether the used direct buffer memory of ZooKeeper reaches 80% of the maximum direct buffer memory specified for ZooKeeper.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **ZooKeeper** > **Configurations** > **All Configurations** > **quorumpeer** > **System**. Check whether **-XX:MaxDirectMemorySize** exists in the **GC_OPTS** parameter.

- If yes, delete **-XX:MaxDirectMemorySize** from **GC_OPTS**. Then go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 6**.

Step 5 Save the configuration and restart the ZooKeeper service.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 6 Check whether the **ALM-13004 ZooKeeper Heap Memory Usage Exceeds the Threshold** exists.

- If yes, handle the alarm by referring to **ALM-13004 ZooKeeper Heap Memory Usage Exceeds the Threshold**.
- If no, go to **Step 7**.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 9 Select **ZooKeeper** in the required cluster from the **Service**.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.88 ALM-13003 GC Duration of the ZooKeeper Process Exceeds the Threshold

Alarm Description

The system checks the garbage collection (GC) duration of the ZooKeeper process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default).

This alarm is cleared when the GC duration is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
13003	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The ZooKeeper process may respond slowly. Services of upper-layer components (such as YARN, Flink, and Spark) may fail.

Possible Causes

The heap memory of the ZooKeeper process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed page, click the drop-down list of **GC Duration of the ZooKeeper Process Exceeds the Threshold**. View the IP address of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **ZooKeeper** > **Instance** > **quorumpeer**. Click the drop-down list in the upper right corner of **Chart**, choose **Customize** > **GC**, select **ZooKeeper GC Duration per Minute**, and click **OK** to check the GC duration statistics of the ZooKeeper process collected every minute.

Step 3 Check whether the GC duration of the ZooKeeper process collected every minute exceeds the threshold (12 seconds by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 Check whether memory leakage occurs in the application.

Step 5 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **ZooKeeper** > **Configurations** > **All Configurations** > **quorumpeer** > **System**. Increase the value of the **GC_OPTS** parameter as required.

NOTE

Generally, **-Xmx** is twice of ZooKeeper data capacity. If the capacity of ZooKeeper reaches 2 GB, set **GC_OPTS** as follows:

-Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=512M -XX:MetaspaceSize=64M -XX:MaxMetaspaceSize=64M -XX:CMSFullGCsBeforeCompaction=1

Step 6 Save the configuration and restart the ZooKeeper service.

NOTE

The service is unavailable during the restart. Upper-layer services that depend on the service are also affected.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **ZooKeeper** for the target cluster.

Step 10 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.89 ALM-13004 ZooKeeper Heap Memory Usage Exceeds the Threshold

Description

The system checks the heap memory usage of the ZooKeeper service every 60 seconds. The alarm is generated when the heap memory usage of a ZooKeeper instance exceeds the threshold (95% of the maximum memory).

The alarm is cleared when the memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13004	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available ZooKeeper heap memory is insufficient, memory overflow may cause service breakdown, and upstream components (such as Yarn, Flink, and Spark) may fail to run.

Possible Causes

The heap memory of the ZooKeeper instance is overused or the heap memory is inappropriately allocated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, On the displayed interface, click the drop-down button of **ZooKeeper Heap Memory Usage Exceeds the Threshold** and confirm the node IP address of the host for which the alarm is generated in the Location Information.

Step 2 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper** > **Instance**, click **quorumpeer** in the **Role** column of the corresponding IP address. Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **CPU and Memory**, and select **ZooKeeper Heap And Direct Buffer Resource Percentage**, click **OK**. Check the heap memory usage.

Step 3 Check whether the used heap memory of ZooKeeper reaches 95% of the maximum heap memory specified for ZooKeeper.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **ZooKeeper** > **Configurations** > **All Configurations** > **quorumpeer** > **System**. Increase the value of **-Xmx** in **GC_OPTS** as required. The details are as follows:

1. On the **Instance** tab, click **quorumpeer** in the **Role** column of the corresponding IP address. Choose **Customize** > **CPU and Memory** in the upper right corner, and select **ZooKeeper Heap And Direct Buffer Resource**, click **OK** to check the heap memory used by ZooKeeper.
2. Change the value of **-Xmx** in the **GC_OPTS** parameter based on the actual heap memory usage. Generally, the value is twice the size of the ZooKeeper data volume. For example, if 2 GB ZooKeeper heap memory is used, the following configurations are recommended: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=512M -XX:MetaspaceSize=64M -XX:MaxMetaspaceSize=64M -XX:CMSFullGCsBeforeCompaction=1

Step 5 Save the configuration and restart the ZooKeeper service.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **ZooKeeper** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.90 ALM-13005 Failed to Set the Quota of Top Directories of ZooKeeper Components

Description

The system sets quotas for each ZooKeeper top-level directory in the **customized.quota** configuration item and components every 5 hours. This alarm is generated when the system fails to set the quota for a directory.

This alarm is cleared when the setting succeeds after a failure.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
13005	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
ServiceDirectory	Specifies the directory for which the alarm is generated.
Trigger Condition	Specifies the cause of the alarm.

Impact on the System

Components can write a large amount of data to the top-level directory of ZooKeeper. As a result, services or services of upstream components (such as Yarn, Flink, and Spark) that depend on the top-level directory are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The quota for the alarm directory is inappropriate.

Procedure

Check whether the quota for the alarm directory is appropriate.

Step 1 Log in to FusionInsight Manager, and choose **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper**. On the displayed page, choose **Configurations** > **All Configurations** > **Quota**. Check whether the directory for which the alarm is reported and its quota exist in the **customized.quota** configuration item.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 Check whether the alarm directory for which the alarm is reported is in the following alarm list.

Table 9-117 Component alarm directory

Component	Alarm Directory
Hbase	/hbase
Hive	/beelinesql
Yarn	/rmstore
Storm	/stormroot
Streaming	/storm
Kafka	/kafka

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 View the component of the alarm directory in the table, open the corresponding service page, and choose **Configurations** > **All Configurations**. On the displayed page, search for **zk.quota** in the upper right corner. The search result is the quota of the alarm directory.

Step 4 Check whether the quota of the alarm directory for which the alarm is reported is appropriate. The quota must be greater than or equal to the actual value, which can be obtained in **Trigger Condition**.

Step 5 Modify the **services.quota** value as prompted and save the configuration.

Step 6 After the time specified by **service.quotas.auto.check.cron.expression**, check whether the alarm is cleared.

The **service.quotas.auto.check.cron.expression** parameter indicates the scheduled expression used by ZooKeeper to set the directory quota. You can choose **Cluster > Services > ZooKeeper > Configurations > All Configurations**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

on FusionInsight Manager and set this parameter. The default value is */5 * * * *, indicating 5 minutes.

- If it is, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **ZooKeeper** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.91 ALM-13006 Znode Number or Capacity Exceeds the Threshold

Description

The system periodically detects the status of secondary Znode in the ZooKeeper service data directory every four hours. This alarm is generated when the number or capacity of secondary Znodes exceeds the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
13006	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceDirectory	Specifies the directory for which the alarm is generated.
Trigger Condition	Specifies the cause of the alarm.

Impact on the System

A large amount of data is written to the ZooKeeper data directory space. As a result, services of upstream components (such as Yarn, Flink, and Spark) that depend on this directory are abnormal. For details, see the alarm location information.

Possible Causes

A large amount of data is written to the ZooKeeper data directory. The threshold is not appropriate.

Procedure

Check whether a large amount of data is written to the directory for which the alarm is generated.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed interface, click the drop-down button of **Znode Number or Capacity Exceeds the Threshold**. Confirm the Znode for which the alarm is generated in Location Information.

Step 2 Log in to FusionInsight Manager, open the ZooKeeper service interface, and select **Resource**. In the table **Used Resources (By Second-Level Znode)**, check whether a large amount of data is written to the top-level Znode for which the alarm is reported.

- If it is, go to **Step 3**.
- If it is not, go to **Step 4**.

Step 3 Log in to the ZooKeeper client and delete the data in the top-level Znode.

Step 4 Log in to FusionInsight Manager and open the ZooKeeper service interface. On the **Resource** page, choose > **By Znode quantity** in **Used Resources (By Second-Level Znode)**. **Threshold Configuration of By Znode quantity** is displayed. Click **Modify** under **Operation**. Increase the threshold by referring to the value of **max.Znode.count** by choosing **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper** > **Configurations > All Configurations > Quota**.

Step 5 In the **Used Resources (By Second-Level Znode)**, choose > **By capacity**. The **Threshold Settings** page of **By Capacity** is displayed. Click **Modify** under **Operation**. Increase the threshold by referring to the value of **max.data.size** by choosing **Cluster >** *Name of the desired cluster* **> Services** > **ZooKeeper** > **Configurations > All Configurations** > **Quota**.

Step 6 Check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is, no further action is required.
- If it is not, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **ZooKeeper** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.92 ALM-13007 Available ZooKeeper Client Connections Are Insufficient

Description

The system periodically detects the number of active processes between the ZooKeeper client and the ZooKeeper server every 60 seconds. This alarm is generated when the number of connections exceeds the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
13007	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
HostName	Specifies the host name for which the alarm is generated.
ClientIP	Specifies the client IP address.
ServerIP	Specifies the server IP address.
Trigger Condition	Specifies the cause of the alarm.

Impact on the System

A large number of processes are connected to ZooKeeper, and the number of ZooKeeper connections is used up. As a result, services of upstream components (such as Yarn, Flink, and Spark) are abnormal.

Possible Causes

A large number of client processes are connected to ZooKeeper. The thresholds are not appropriate.

Procedure

Check whether there are a large number of client processes connected to ZooKeeper.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed interface, click the drop-down button of **Available ZooKeeper Client Connections Are Insufficient**. Confirm the node IP address of the host for which the alarm is generated in the Location Information.

Step 2 Open the ZooKeeper service interface, click **Resource** to enter the **Resource** page, and check whether the number of connections of the client with the IP address specified by **Number of Connections (By Client IP Address)** is large.

- If it is, go to **Step 3**.
- If it is not, go to **Step 4**.

Step 3 Check whether connection leakage occurs on the client process.

Step 4 Click in the **Number of Connections (by Client IP Address)** to enter the **Thresholds** page, and click **Modify** under **Operation**. Increase the threshold by referring to the value of **maxClientCnxns** by choosing **Cluster** > *Name of the desired cluster* > **Services** > **ZooKeeper** > **Configurations** > **All Configurations** > **quorumpeer**.

Step 5 Check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select **ZooKeeper** in the required cluster from the **Service**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.93 ALM-13008 ZooKeeper Znode Usage Exceeds the Threshold

Alarm Description

The system checks the level-2 Znode status in the ZooKeeper data directory every hour. This alarm is generated when the system detects that the level-2 Znode usage exceeds the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
13008	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
ServiceDirectory	Specifies the directory for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Trigger Condition	Specifies the cause of the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

When a large amount of data is written to the ZooKeeper data directory space, ZooKeeper cannot provide services for external systems. As a result, services of upstream components (such as Yarn, Flink, and Spark) that depend on the alarm directory are abnormal.

Possible Causes

- A large amount of data is written to the ZooKeeper data directory.
- The user-defined threshold is inappropriate.

Handling Procedure

Check whether a large amount of data is written into the directory for which the alarm is generated.

Step 1 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **ZooKeeper**, and click **Resource**. Click **By Znode quantity** in **Used Resources (By Second-Level Znode)**, and check whether a large amount of data is written to the top Znode.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Log in to FusionInsight Manager, choose **O&M > Alarm > Alarms**, select **Location** from the drop-down list box next to **ALM-13008 ZooKeeper Znode Quantity Usage Exceeds Threshold**, and obtain the Znode path in **ServiceDirectory**.

Step 3 Log in to the ZooKeeper client as a cluster user and delete unnecessary data from the Znode corresponding to the alarm.

Step 4 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **ZooKeeper** > **Configurations** > **All Configurations**, and search for **max.znode.count**, which is the maximum number of ZooKeeper directories. The alarm threshold is 80% of this parameter. Increase the value of this parameter, click **Save**, and restart the service for the configuration to take effect.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select **ZooKeeper** in the required cluster from the **Service**.

Step 8 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.94 ALM-13009 ZooKeeper ZNode Capacity Usage Exceeds the Threshold

Alarm Description

The system checks the level-2 znode status in the ZooKeeper data directory every hour. This alarm is generated when the system detects that the total capacity usage exceeds the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
13009	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
ServiceDirectory	Specifies the directory for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

ZooKeeper cannot provide services for external systems, and the services of upstream components (such as YARN, Flink, and Spark) that depend on the alarm directory are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- A large volume of data has been written to the ZooKeeper data directory.
- The threshold is improperly defined.

Handling Procedure

Check whether a large volume of data is written to the alarm directory.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. Click the drop-down list in the row containing **ALM-13009 ZooKeeper ZNode Capacity Usage Exceeds the Threshold**, and find the ZNode for which the alarm is generated in the **Location** area.

Step 2 Choose **Cluster** > **Services** > **ZooKeeper**. On the page that is displayed, click the **Resource** tab. In the **Used Resources (By Second-Level ZNode)** area, click **By capacity** and check whether a large amount of data is written to the top-level ZNode directory.

- If yes, record the directory to which a large amount of data is written and go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Check whether data in the directory can be deleted.

NOTICE

Deleting data from ZooKeeper is a high-risk operation. Exercise caution when performing this operation.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Log in to the ZooKeeper client and delete unnecessary data from the directory to which a large amount of data is written.

1. Log in to the ZooKeeper client installation directory, for example, **/opt/client**, and configure environment variables.

   cd /opt/client

   source bigdata_env

2. Run the following command to authenticate the user (skip this step for a cluster in normal mode):

   kinit *Component service user*

3. Run the following command to log in to the client tool:

   zkCli.sh -server <*Service IP address of the node where any ZooKeeper instance resides*>:<*Client port*>

4. Run the following command to delete unnecessary data:

   delete *Path of the file to be deleted*

Step 5 Log in to FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **ZooKeeper** > **Configurations** > **All Configurations**, and search for **max.data.size**. The value of **max.data.size** is the maximum

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

capacity quota of the ZooKeeper directory. The unit is byte. Search for the **GC_OPTS** configuration item and check the value of **Xmx**.

Step 6 Compare the values of **max.data.size** and **Xmx*0.65**. The threshold is the smaller value multiplied by 80%. You can change the values of **max.data.size** and **Xmx*0.65** to increase the threshold.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **ZooKeeper** for the target cluster.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.95 ALM-13010 Znode Usage of a Directory with Quota Configured Exceeds the Threshold

Alarm Description

The system checks the Znode usage of all service directories with quota configured every hour. This alarm is generated when the system detects that the level-2 Znode usage exceeds the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
13010	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
ServiceDirectory	Specifies the directory for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Trigger Condition	Specifies the cause of the alarm.

Impact on the System

When a large amount of data is written to the ZooKeeper data directory space, ZooKeeper cannot provide services for external systems. As a result, services of upper-layer components (such as Yarn, Flink, and Spark) that depend on the alarm directory are abnormal.

Possible Causes

- A large amount of data is written to the ZooKeeper data directory.
- The user-defined threshold is inappropriate.

Handling Procedure

Check whether a large amount of data is written into the directory for which the alarm is generated.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm > Alarms**. Confirm the Znode for which the alarm is generated in **Location** of this alarm.

Step 2 Choose **Cluster** > **Services** > **ZooKeeper** and click **Resource**. In **Used Resources (By Second-Level Znode)**, check whether a large amount of data is written into the top Znode.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Log in to FusionInsight Manager, choose **O&M > Alarm > Alarms**, click the arrow next to **ALM-13010 Znode Usage of a Directory with Quota Configured Exceeds the Threshold**, and obtain the Znode path from **ServiceDirectory** in **Location**.

Step 4 Log in to the ZooKeeper client as a cluster user and delete unwanted data in the Znode for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Log in to FusionInsight Manager, and choose **Cluster** > **Services** > *Component of the top Znode for which the alarm is generated*. Choose **Configurations** > **All Configurations**, search for **zk.quota.number**, increase its value, click **Save**.

NOTICE

If the Component of the top Znode for which the alarm is generated is ClickHouse, change the value of **clickhouse.zookeeper.quota.node.count**.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **ZooKeeper** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.96 ALM-14000 HDFS Service Unavailable

Description

The system checks the NameService service status every 60 seconds. This alarm is generated when all the NameService services are abnormal and the system considers that the HDFS service is unavailable.

This alarm is cleared when at least one NameService service is normal and the system considers that the HDFS service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14000	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

HDFS fails to provide services for HDFS service-based upper-layer components, such as HBase and MapReduce. As a result, users cannot read or write files.

Possible Causes

- The ZooKeeper service is abnormal.
- All NameService services are abnormal.
- The number of service requests is too large, and the HDFS health check fails to read and write files.
- The health check fails due to HDFS FullGC.

Procedure

Check the ZooKeeper service status.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. On the Alarm page, check whether **ALM-13000 ZooKeeper Service Unavailable** is reported.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Rectify the ZooKeeper service fault by following the instructions in **ALM-13000 ZooKeeper Service Unavailable**. Then check whether the running status of the ZooKeeper service is normal.

- If yes, go to **Step 3**.
- If no, go to **Step 13**.

Step 3 On the **O&M > Alarm > Alarms** page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Handle the NameService service exception alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. On the Alarms page, check whether **ALM-14010 NameService Service Unavailable** is reported.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Handle the abnormal NameService services by referring to the handling method provided in **ALM-14010 NameService Service Is Abnormal**. Then, check whether the alarm is cleared.

- If yes, go to **Step 6**.
- If no, go to **Step 13**.

Step 6 On the **O&M > Alarm > Alarms** page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check whether the HDFS health check fails to read or write files due to a large number of service requests.

Step 7 On FusionInsight Manager, choose **O&M > Alarm > Alarms**, and check whether **ALM-14021 NameNode Average RPC Processing Time Exceeds the Threshold** or **ALM-14022 NameNode Average RPC Queuing Time Exceeds the Threshold** is generated.

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

Step 8 Handle the abnormal NameService services by following the instructions provided in **ALM-14021 NameNode Average RPC Processing Time Exceeds the Threshold** and **ALM-14022 NameNode Average RPC Queuing Time Exceeds the Threshold**, and check whether the alarm is cleared for each NameService.

- If yes, go to **Step 9**.
- If no, go to **Step 13**.

Step 9 On the **O&M > Alarm > Alarms** page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether the health check fails due to HDFS FullGC.

Step 10 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. On the Alarms page, check whether **ALM-14014 NameNode GC Time Exceeds the Threshold** is reported.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Handle the abnormal NameService services by referring to the handling method provided in **ALM-14014 NameNode GC Time Exceeds the Threshold**. Then check whether the alarm is cleared for each NameService service.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 On the **O&M > Alarm > Alarms** page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 14 Select the following nodes in the required cluster from the **Service**:

- ZooKeeper
- HDFS

Step 15 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.97 ALM-14001 HDFS Disk Usage Exceeds the Threshold

Alarm Description

The system checks the HDFS disk usage every 30 seconds and compares the actual HDFS disk usage with the threshold. The HDFS disk usage metric has a default threshold. This alarm is generated when the HDFS disk usage exceeds the threshold.

You can choose **O&M** > **Alarm** > **Thresholds** > **HDFS** and change the threshold.

When **Trigger Count** is **1**, this alarm is cleared when the value of the disk usage of HDFS cluster indicator is less than or equal to the threshold. When **Trigger Count** is greater than **1**, this alarm is cleared when the value of the disk usage of HDFS cluster indicator is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14001	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NameServiceName	Specifies the NameService for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Writing Hadoop distributed file system (HDFS) data is affected.

Possible Causes

The disk space configured for the HDFS cluster is insufficient.

Handling Procedure

Check the disk capacity and delete unnecessary files.

Step 1 On the homepage of FusionInsight Manager, choose **Cluster** > **Services** > **HDFS**.

Step 2 In the **Chart** area, click **Real-time**, and view the value of the monitoring item **Percentage of HDFS Capacity** to check whether the HDFS disk usage exceeds the threshold (**80%** by default). If the monitoring item is not displayed, click the drop-down list in the upper right corner of the **Chart** area, choose **Customize** > **Disk**, and select **Percentage of HDFS Capacity**.

- If yes, go to **Step 3**.
- If no, go to **Step 11**.

Step 3 In the **Basic Information** area, click the **NameNode(Active)** of the failure NameService and the HDFS WebUI page is displayed.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On the HDFS web user interface (WebUI), click **Datanodes** tab. In the **Block pool used** column, view the disk usage of all DataNodes to check whether the disk usage of any DataNode exceeds the threshold.

- If yes, go to **Step 6**.
- If no, go to **Step 11**.

Step 5 Log in to the node as user **root** where the cluster client is installed.

Step 6 Run the following command to go to the client installation directory and complete authentication:

```
cd /opt/client
```

Load environment variables:

```
source bigdata_env
```

If Kerberos authentication is enabled for the cluster, run the following command to complete authentication as the cluster user that has the operation permissions for HDFS.

You need to create related users on Manager in advance or contact the cluster administrator to obtain the users. For details, see **Creating an MRS Cluster User**.

```
kinit hdfsuser
```

Step 7 Run the following command to delete unnecessary HDFS files:

1. Run the following command to check files or sub-directories in the corresponding folder:
   hdfs dfs -ls *Folder name*
2. Run the following command to query the HDFS file directory size:
   hdfs dfs -du *HDFS file directory*

   For example, if you run the above command to query the size of files or sub-directories in the **/tmp** folder, the following command output is returned, indicating that the **/tmp/hadoop-yarn** directory occupies 622,222,017 bytes:

   ```
   0          0          /tmp/.testHDFS
   207407339  622222017  /tmp/hadoop-yarn
   0          0          /tmp/hive
   0          0          /tmp/hive-scratch
   0          0          /tmp/logs
   ...
   ```

3. Run the following command to delete confirmed unnecessary HDFS files:
   hdfs dfs -rm -r *HDFS file or directory path*

WARNING

Deleting files is a high-risk operation. Ensure that the files to be deleted are unnecessary files to prevent data loss.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Expand the system.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Expand the capacity of the cluster node where HDFS is deployed. For details, see **Scaling Out an MRS Cluster**.

Step 10 Then, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 12 Select the following nodes in the required cluster from the **Service**:

- ZooKeeper
- HDFS

Step 13 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 14 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.98 ALM-14002 DataNode Disk Usage Exceeds the Threshold

Alarm Description

The system checks the DataNode disk usage every 30 seconds and compares the actual disk usage with the threshold. A default threshold range is provided for the DataNode disk usage. This alarm is generated when the DataNode disk usage exceeds the threshold.

You can choose **O&M** > **Alarm** > **Thresholds** > **HDFS** and change the threshold.

If **Trigger Count** is **1**, this alarm is cleared when the DataNode disk usage is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the DataNode disk usage is less than or equal to 80% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14002	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Insufficient disk space will impact data write to HDFS.

Possible Causes

- The disk space configured for the HDFS cluster is insufficient.
- Data skew occurs among DataNodes.

Handling Procedure

Check whether the cluster disk capacity is insufficient.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether the **ALM-14001 HDFS Disk Usage Exceeds the Threshold** alarm exists.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Handle the alarm by following the instructions in **ALM-14001 HDFS Disk Usage Exceeds the Threshold** and check whether the alarm is cleared.

- If yes, go to **Step 3**.
- If no, go to **Step 11**.

Step 3 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the balance status of DataNodes.

Step 4 On FusionInsight Manager, choose **Hosts**. Check whether the number of DataNodes on each rack is almost the same. If the difference is large, adjust the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

racks to which DataNodes belong to ensure that the number of DataNodes on each rack is almost the same. Restart the HDFS service for the settings to take effect.

NOTE

The service is unavailable during the restart, and upper-layer services that depend on the service will also be affected.

Step 5 Choose **Cluster** > **Services** > **HDFS**.

Step 6 In the **Basic Information** area, click **NameNode(Active)**. The HDFS web UI is displayed.

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 7 In the **Summary** area of the HDFS web UI, check whether the value of **Max** is 10% greater than that of **Median** in **DataNodes usages**.

- If yes, go to **Step 8**.
- If no, go to **Step 11**.

Step 8 Balance skewed data in the cluster.

Log in to the node where the MRS cluster client is installed, as user **root**.

1. If Kerberos authentication is not enabled for the MRS cluster, run the following command to switch to user **omm**:
   ```
   su - omm
   ```
2. Run the following command to go to the client installation directory and load environment variables:
   ```
   cd Client installation directory
   source bigdata_env
   ```
3. If Kerberos authentication is enabled for the cluster, complete security authentication:
   ```
   kinit Username of the service user who has the HDFS operation permissions
   ```
   Enter the required password.

Step 9 Run the following command to balance data distribution:

```
hdfs balancer -threshold 10
```

-threshold < Percentage >: specifies the balancing threshold (10% by default).

When the difference between the node storage usage and the average cluster storage usage exceeds the threshold, data migration is triggered.

Step 10 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect the fault information.

Step 11 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 13 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.99 ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold

Alarm Description

The system checks the lost HDFS blocks every 30 seconds and compares the actual lost blocks with the threshold. This alarm is generated when the number of lost HDFS blocks usage exceeds the threshold.

You can choose **O&M** > **Alarm** > **Thresholds** > **HDFS** > **Lost Blocks** and change the threshold.

If **Trigger Count** is **1**, this alarm is cleared when the value of lost HDFS blocks is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the value of lost HDFS blocks is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14003	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
NameServiceName	Specifies the NameService for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Data stored in HDFS is lost. HDFS may enter the security mode and cannot provide write services. Lost block data cannot be restored.

Possible Causes

- The DataNode instance is abnormal.
- HDFS data is deleted.

Handling Procedure

Check the DataNode instance.

Step 1 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS** > **Instances**.

Step 2 Check whether the **Running Status** of all DataNode instance is **Normal**.

- If yes, go to **Step 11**.
- If no, go to **Step 3**.

Step 3 Restart the DataNode instance and check whether the DataNode instance restarts successfully.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Delete the damaged file.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS** > **NameNode(Active)**. On the WebUI page of the HDFS, view the information about lost blocks.

- If a block is lost, an error message is displayed on the HDFS web UI.
- For clusters with Kerberos authentication enabled, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 6 The user checks whether the file containing the lost data block is useful.

NOTE

Files generated in directories **/mr-history**, **/tmp/hadoop-yarn**, and **/tmp/logs** during MapReduce task execution are unnecessary.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 The user checks whether the file containing the lost data block is backed up.

- If yes, go to **Step 8**.
- If no, go to **Step 11**.

Step 8 Log in to the node where the HDFS cluster client is installed, as user **root**. The user password is defined by the user before the installation. Contact the MRS cluster administrator to obtain the password.

- For clusters with Kerberos authentication enabled:
  ```
  cd Client installation directory
  source bigdata_env
  kinit hdfs
  ```
- For clusters with Kerberos authentication disabled:
  ```
  su - omm
  cd Client installation directory
  source bigdata_env
  ```

Step 9 Run the following command to delete a lost file. If the file where the lost block is located is a useful file, you need to write the file again to restore file data.

```
hdfs fsck / -delete
```

WARNING

Deleting a file or folder is a high-risk operation. Ensure that the file or folder is no longer required before performing this operation.

Step 10 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect the fault information.

Step 11 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 12 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.100 ALM-14006 Number of HDFS Files Exceeds the Threshold

Alarm Description

The system periodically checks the number of HDFS files every 30 seconds and compares the number of HDFS files with the threshold. This alarm is generated when the system detects that the number of HDFS files exceeds the threshold.

If **Trigger Count** is **1**, this alarm is cleared when the number of HDFS files is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the number of HDFS files is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14006	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
NameServiceName	Specifies the NameService for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If there are too many HDFS files, the HDFS system may respond slowly or the disk space may be used up.

Possible Causes

The number of HDFS files exceeds the threshold.

Handling Procedure

Check the number of files in the system.

Step 1 On FusionInsight Manager, check the number of HDFS files. Specifically, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS**. Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **File and Block**, and select **HDFS File** and **Total Blocks**.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Configurations** > **All Configurations**, and search for the **GC_OPTS** parameter under **NameNode**.

Step 3 Configure the threshold of the number of configuration file objects. Specifically, change the value of **Xmx** (GB) in the **GC_OPTS** parameter. The threshold (specified by y) is calculated as follows: y = 0.2007 x Xmx - 0.6312, where x indicates the memory capacity Xmx (GB) and y indicates the number of files (unit: kW). Adjust the memory size as required.

Step 4 Check that the value of **GC_PROFILE** is **custom** so that the **GC_OPTS** configuration can take effect. Click **Save**, click **More**, and select **Restart Service** to restart the service.

NOTE

The service is unavailable during the restart. Upper-layer services that depend on the service are also affected.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether needless files exist in the system.

Step 6 Log in to the HDFS client as user **root**. Run **cd** to switch to the client installation directory, and run **source bigdata_env** to configure the environment variables.

If the cluster uses the security mode, perform security authentication.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Run the **kinit hdfs** command and enter the password as prompted. Obtain the password from the MRS cluster administrator.

Step 7 Run **hdfs dfs -ls** *file or directory* to check whether the files in the directory can be deleted.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Run the **hdfs dfs -rm -r** *file or directory path* command. After deleting unnecessary files, wait until the files are retained in the recycle bin for a period longer than the value of **fs.trash.interval** on the NameNode. Then check whether the alarm is cleared.

NOTE

Deleting a file or folder is a high-risk operation. Ensure that the file or folder is no longer required before performing this operation.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 10 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

Configuration rules of the NameNode JVM parameter

Default value of the NameNode JVM parameter **GC_OPTS**:

-Xms2G -Xmx4G -XX:NewSize=128M -XX:MaxNewSize=256M -XX:MetaspaceSize=128M -XX:MaxMetaspaceSize=128M -XX:+UseConcMarkSweepGC -XX:+CMSParallelRemarkEnabled -XX:CMSInitiatingOccupancyFraction=65 -XX:+PrintGCDetails -Dsun.rmi.dgc.client.gcInterval=0x7FFFFFFFFFFFFE -Dsun.rmi.dgc.server.gcInterval=0x7FFFFFFFFFFFFE -XX:-OmitStackTraceInFastThrow -XX:+PrintGCDateStamps -XX:+UseGCLogFileRotation -XX:NumberOfGCLogFiles=10 -XX:GCLogFileSize=1M -Djdk.tls.ephemeralDHKeySize=3072 -Djdk.tls.rejectClientInitiatedRenegotiation=true -Djava.io.tmpdir=${Bigdata_tmp_dir}

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The number of NameNode files is proportional to the used memory size of the NameNode. When file objects change, you need to change **-Xms2G -Xmx4G -XX:NewSize=128M -XX:MaxNewSize=256M** in the default value. The following table lists the reference values.

Table 9-118 NameNode JVM configuration

Number of File Objects	Reference Value
10,000,000	-Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
20,000,000	-Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G
50,000,000	-Xms32G -Xmx32G -XX:NewSize=3G -XX:MaxNewSize=3G
100,000,000	-Xms64G -Xmx64G -XX:NewSize=6G -XX:MaxNewSize=6G
200,000,000	-Xms96G -Xmx96G -XX:NewSize=9G -XX:MaxNewSize=9G
300,000,000	-Xms164G -Xmx164G -XX:NewSize=12G -XX:MaxNewSize=12G

9.11.101 ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold

Description

The system checks the HDFS NameNode Heap Memory usage every 30 seconds and compares the actual Heap memory usage with the threshold. The HDFS NameNode Heap Memory usage has a default threshold. This alarm is generated when the HDFS NameNode Heap Memory usage exceeds the threshold.

You can change the threshold in **O&M** > **Alarm > Thresholds** > *Name of the desired cluster* **> HDFS**.

When the **Trigger Count** is 1, this alarm is cleared when the HDFS NameNode Heap memory usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the HDFS NameNode Heap memory usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14007	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The HDFS NameNode Heap Memory usage is too high, which affects the data read/write performance of the HDFS.

Possible Causes

The HDFS NameNode Heap Memory is insufficient.

Procedure

Delete unnecessary files.

Step 1 Log in to the HDFS client as user **root**. Run **cd** to switch to the client installation directory, and run **source bigdata_env**.

If the cluster uses the security mode, perform security authentication.

Run the **kinit hdfs** command and enter the password as prompted. Obtain the password from the administrator.

Step 2 Run the **hdfs dfs -rm -r** *file or directory* command to delete unnecessary files.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the NameNode JVM memory usage and configuration.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS**.

Step 5 In the **Basic Information** area, click **NameNode(Active)** to go to the HDFS WebUI.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 6 On the HDFS WebUI, click the **Overview** tab. In **Summary**, check the numbers of files, directories, and blocks in the HDFS.

Step 7 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Configurations** > **All Configurations**. In **Search**, enter **GC_OPTS** to check the **GC_OPTS** memory parameter of **HDFS->NameNode**.

Adjust the configuration in the system.

Step 8 Check whether the memory is configured properly based on the number of files in **Step 6** and the NameNode Heap Memory parameters in **Step 7**.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

NOTE

The recommended mapping between the number of HDFS file objects (filesystem objects = files + blocks) and the JVM parameters configured for NameNode is as follows:

- If the number of file objects reaches 10,000,000, you are advised to set the JVM parameters as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
- If the number of file objects reaches 20,000,000, you are advised to set the JVM parameters as follows: -Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G
- If the number of file objects reaches 50,000,000, you are advised to set the JVM parameters as follows: -Xms32G -Xmx32G -XX:NewSize=3G -XX:MaxNewSize=3G
- If the number of file objects reaches 100,000,000, you are advised to set the JVM parameters as follows: -Xms64G -Xmx64G -XX:NewSize=6G -XX:MaxNewSize=6G
- If the number of file objects reaches 200,000,000, you are advised to set the JVM parameters as follows: -Xms96G -Xmx96G -XX:NewSize=9G -XX:MaxNewSize=9G
- If the number of file objects reaches 300,000,000, you are advised to set the JVM parameters as follows: -Xms164G -Xmx164G -XX:NewSize=12G -XX:MaxNewSize=12G

Step 9 Modify the heap memory parameters of the NameNode based on the mapping between the number of file objects and the memory. Click **Save** and choose **Dashboard** > **More** > **Restart Service**.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 10 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 12 Select the following nodes in the required cluster from the **Service**:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- ZooKeeper
- HDFS

Step 13 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.102 ALM-14008 DataNode Heap Memory Usage Exceeds the Threshold

Description

The system checks the HDFS DataNode Heap Memory usage every 30 seconds and compares the actual Heap Memory usage with the threshold. The HDFS DataNode Heap Memory usage has a default threshold. This alarm is generated when the HDFS DataNode Heap Memory usage exceeds the threshold.

You can change the threshold in **O&M** > **Alarm > Thresholds** > *Name of the desired cluster* **> HDFS**.

When the **Trigger Count** is 1, this alarm is cleared when the HDFS DataNode Heap Memory usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the HDFS DataNode Heap Memory usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14008	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The HDFS DataNode Heap Memory usage is too high, which affects the data read/write performance of the HDFS.

Possible Causes

The HDFS DataNode Heap Memory is insufficient.

Procedure

Delete unnecessary files.

Step 1 Log in to the HDFS client as user **root**. Run **cd** to switch to the client installation directory, and run **source bigdata_env**.

If the cluster uses the security mode, perform security authentication.

Run the **kinit hdfs** command and enter the password as prompted. Obtain the password from the administrator.

Step 2 Run the **hdfs dfs -rm -r** *file or directory* command to delete unnecessary files.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the DataNode JVM memory usage and configuration.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS**.

Step 5 In the **Basic Information** area, click **NameNode(Active)** to go to the HDFS WebUI.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 6 On the HDFS WebUI, click the **DataNodes** tab, and check the number of blocks of all DataNodes related to the alarm.

Step 7 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Configurations** > **All Configurations**. In **Search**, enter **GC_OPTS** to check the GC_OPTS memory parameter of **HDFS->DataNode**.

Adjust the configuration in the system.

Step 8 Check whether the memory is configured properly based on the number of blocks in **Step 6** and the DataNode Heap Memory parameters in **Step 7**.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

NOTE

The mapping between the average number of blocks of a DataNode instance and the DataNode memory is as follows:

- If the average number of blocks of a DataNode instance reaches 2,000,000, the reference values of the JVM parameters of the DataNode are as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
- If the average number of blocks of a DataNode instance reaches 5,000,000, the reference values of the JVM parameters of the DataNode are as follows: -Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G

Step 9 Modify the heap memory parameters of the DataNode based on the mapping between the number of blocks and the memory. Click **Save** and choose **Dashboard** > **More** > **Restart Service**.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 10 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 12 Select **HDFS** in the required cluster from the **Service**.

Step 13 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.103 ALM-14009 Number of Dead DataNodes Exceeds the Threshold

Description

The system periodically detects the number of dead DataNodes in the HDFS cluster every 30 seconds, and compares the number with the threshold. The number of DataNodes in the Dead state has a default threshold. This alarm is generated when the number exceeds the threshold.

You can change the threshold in **O&M** > **Alarm > Thresholds** > *Name of the desired cluster* **> HDFS**.

When the **Trigger Count** is 1, this alarm is cleared when the number of Dead DataNodes is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the number of Dead DataNodes is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14009	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NameServiceName	Specifies the NameService for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

DataNodes that are in the Dead state cannot provide HDFS services. As a result, users cannot read or write files.

Possible Causes

- DataNodes are faulty or overloaded.
- The network between the NameNode and the DataNode is disconnected or busy.
- NameNodes are overloaded.
- The NameNodes are not restarted after the DataNode is deleted.

Procedure

Check whether DataNodes are faulty.

Step 1 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS**. The **HDFS Status** page is displayed.

Step 2 In the **Basic Information** area, click **NameNode(Active)** to go to the HDFS WebUI.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 3 On the HDFS WebUI, click the **Datanodes** tab. In the **In operation** area, click **Filter** to check whether **down** is in the drop-down list.

- If yes, select **down**, record the information about the filtered DataNodes, and go to **Step 4**.
- If no, go to **Step 8**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Instance** to check whether recorded DataNodes exist in the instance list.

- If all recorded DataNodes exist, go to **Step 5**.
- If no recorded DataNodes exists, go to **Step 6**.
- If some of the recorded DataNodes exist, go to **Step 7**.

Step 5 Locate the DataNode instance, click **More** > **Restart Instance** to restart it and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform this operation during off-peak hours.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 6 Select all NameNode instances, choose **More** > **Instance Rolling Restart** to restart them and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Step 7 Select all NameNode instances, choose **More** > **Instance Rolling Restart** to restart them. Locate the DataNode instance, click **More** > **Restart Instance** to restart it and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check the status of the network between the NameNode and the DataNode.

Step 8 Log in to the faulty DataNode on the management page as user **root**, and run the **ping** *IP address of the NameNode* command to check whether the network between the DataNode and the NameNode is abnormal.

On the FusionInsight Manager page, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Instance**. In the instance list, view the service plane IP address of the faulty DataNode.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Rectify the network fault, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether the DataNode is overloaded.

Step 10 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** and check whether the alarm **ALM-14008 HDFS DataNode Memory Usage Exceeds the Threshold** exists.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Rectify the fault by following the handling procedure of **ALM-14008 DataNode Heap Memory Usage Exceeds the Threshold**. Then, check whether the alarm is cleared.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 Check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 13**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the NameNode is overloaded.

Step 13 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and check whether the alarm **ALM-14007 HDFS NameNode Memory Usage Exceeds the Threshold** exists.

- If yes, go to **Step 14**.
- If no, go to **Step 16**.

Step 14 Handle the alarm by referring to the handling procedure in **ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold** and check whether the alarm is cleared.

- If yes, go to **Step 15**.
- If no, go to **Step 16**.

Step 15 Check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 16**.

Collect fault information.

Step 16 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 17 Select **HDFS** in the required cluster from the **Service**.

Step 18 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 19 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.104 ALM-14010 NameService Service Is Abnormal

Alarm Description

The system checks the NameService service status every 180 seconds. This alarm is generated when the NameService service is unavailable.

This alarm is cleared when the NameService service recovers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14010	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
NameServiceName	Specifies the NameService for which the alarm was generated.

Impact on the System

HDFS fails to provide services for upper-layer components based on the NameService service, such as HBase and MapReduce. As a result, users cannot read or write files.

Possible Causes

- The KrbServer service is abnormal.
- The JournalNode is faulty.
- The DataNode is faulty.
- The disk capacity is insufficient.
- The NameNode enters safe mode.

Handling Procedure

Check the KrbServer service status.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services**.

Step 2 Check whether the KrbServer service exists.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click **KrbServer**.

Step 4 Click **Instances**. On the KrbServer management page, select the faulty instance, and choose **More** > **Restart Instance**. Check whether the instance successfully restarts.

- If yes, go to **Step 5**.
- If no, go to **Step 24**.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check the JournalNode instance status.

Step 6 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services**.

Step 7 Choose **HDFS** > **Instances**.

Step 8 Check whether the **Running Status** of the JournalNode is **Normal**.

- If yes, go to **Step 11**.
- If no, go to **Step 9**.

Step 9 Select the faulty JournalNode, and choose **More** > **Restart Instance**. Check whether the JournalNode successfully restarts.

NOTE

If the number of restarted JournalNode instances exceeds one third of the total number of JournalNode instances, the HDFS service may be faulty.

- If yes, go to **Step 10**.
- If no, go to **Step 24**.

Step 10 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check the DataNode instance status.

Step 11 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS**.

Step 12 Click **Instances** and check whether **Running Status** of all DataNodes is **Normal**.

- If yes, go to **Step 15**.
- If no, go to **Step 13**.

Step 13 Click **Instances**. On the DataNode management page, select the faulty instance, and choose **More** > **Restart Instance**. Check whether the DataNode successfully restarts.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

- If yes, go to **Step 14**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 15**.

Step 14 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Check disk status.

Step 15 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Host**.

Step 16 In the **Disk** column, check whether the disk space is insufficient.

- If yes, go to **Step 17**.
- If no, go to **Step 19**.

Step 17 Expand the disk capacity.

Step 18 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 19**.

Check whether NameNode is in the safe mode.

Step 19 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS**. Click **NameNode(Active)** of the abnormal NameService. The NameNode web UI is displayed.

NOTE

By default, the admin user does not have the management rights of other components. If the page cannot be opened or the content is not completely displayed due to insufficient permission when you access the native page of a component, you can manually create a user with the management rights of the corresponding component to log in to the component.

Step 20 On the NameNode web UI, check whether "Safe mode is ON." is displayed.

Information behind **Safe mode is ON** is alarm information and is displayed based actual conditions.

- If yes, go to **Step 21**.
- If no, go to **Step 24**.

Step 21 Log in to the client as user **root**. Run the **cd** command to go to the client installation directory and run the **source bigdata_env** command. If the cluster uses the security mode, perform security authentication. Run the **kinit hdfs** command and enter the password as prompted. The password can be obtained from the MRS cluster administrator. If the cluster uses the non-security mode, log in as user **omm** and run the command. Ensure that user **omm** has the client execution permission.

Step 22 Run **hdfs dfsadmin -safemode leave**.

Step 23 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 24**.

Collect the fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 24 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 25 In the **Service** area, select the following nodes of the desired cluster.

- ZooKeeper
- HDFS

Step 26 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 27 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.105 ALM-14011 DataNode Data Directory Is Not Configured Properly

Description

The DataNode parameter **dfs.datanode.data.dir** specifies DataNode data directories. This alarm is generated when a configured data directory cannot be created, a data directory uses the same disk as other critical directories in the system, or multiple directories use the same disk immediately.

This alarm is cleared when the DataNode data directory is configured properly and this DataNode for which the alarm is generated is restarted.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14011	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the DataNode data directory is mounted to the root directory or a critical directory, the disk space of the root directory or critical directory will be used up after long time running and the system will be faulty.

If the DataNode data directory is not configured properly, HDFS performance will deteriorate.

Possible Causes

- The DataNode data directory fails to be created.
- The DataNode data directory uses the same disk with critical directories, such as **/** or **/boot**.
- Multiple directories in the DataNode data directory use the same disk.

Procedure

Check the alarm cause and information about the DataNode for which the alarm is generated.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. In the alarm list, click the alarm.

Step 2 In **HostName** of **Location**, obtain the host name of the DataNode for which the alarm is generated.

Delete directories that do not comply with the disk plan from the DataNode data directory.

Step 3 Choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Instance**. In the instance list, click the DataNode instance on the node for which the alarm is generated.

Step 4 Click **Instance Configurations** and view the value of the DataNode parameter **dfs.datanode.data.dir**.

Step 5 Check whether all DataNode data directories are consistent with the disk plan.

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 Modify the DataNode parameter **dfs.datanode.data.dir** and delete the incorrect directories.

Step 7 Choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Instance** and restart the DataNode instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform this operation during off-peak hours.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Log in to the DataNode for which the alarm is generated as user **root**.

- If the alarm cause is "The DataNode data directory fails to be created", go to **Step 10**.
- If the alarm cause is "The DataNode data directory uses the same disk with critical directories, such **/** or **/boot**", go to **Step 17**.
- If the alarm cause is "Multiple directories in the DataNode data directory uses the same disk", go to **Step 21**.

Check whether the DataNode data directory fails to be created.

Step 10 Run the **su - omm** command to switch to user **omm**.

Step 11 Run the **ls** command to check whether the directories exist in the DataNode data directory.

- If yes, go to **Step 26**.
- If no, go to **Step 12**.

Step 12 Run the **mkdir** *data directory* command to create the directory and check whether the directory can be successfully created.

- If yes, go to **Step 24**.
- If no, go to **Step 13**.

Step 13 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** to check whether alarm **ALM-12017 Insufficient Disk Capacity** exists.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Adjust the disk capacity and check whether alarm **ALM-12017 Insufficient Disk Capacity** is cleared. For details, see **ALM-12017 Insufficient Disk Capacity**.

- If yes, go to **Step 12**.
- If no, go to **Step 15**.

Step 15 Check whether user **omm** has the **rwx** or **x** permission of all the upper-layer directories of the directory. (For example, for **/tmp/abc/**, user **omm** has the **x** permission for directory **tmp** and the **rwx** permission for directory **abc**.)

- If yes, go to **Step 24**.
- If no, go to **Step 16**.

Step 16 Run the **chmod u+rwx** *path* or **chmod u+x** *path* command as user **root** to assign the **rwx** or **x** permission of these directories to user **omm**. Then go to **Step 12**.

Check whether the DataNode data directory uses the same disk as other critical directories in the system.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 17 Run the **df** command to obtain the disk mounting information of each directory in the DataNode data directory.

Step 18 Check whether the directories mounted to the disk are critical directories, such as **/** or **/boot**.

- If yes, go to **Step 19**.
- If no, go to **Step 24**.

Step 19 Change the value of the DataNode parameter **dfs.datanode.data.dir** and delete the directories that use the same disk as critical directories.

Step 20 Go to **Step 24**.

Check whether multiple directories in the DataNode data directory use the same disk.

Step 21 Run the **df** command to obtain the disk mounting information of each directory in the DataNode data directory. Record the mounted directory in the command output.

Step 22 Modify the DataNode node parameters **dfs.datanode.data.dir** to reserve only one directory among the directories that mounted to the same disk directory.

Step 23 Go to **Step 24**.

Restart the DataNode and check whether the alarm is cleared.

Step 24 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Instance** and restart the DataNode instance.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform this operation during off-peak hours.

Step 25 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 26**.

Collect fault information.

Step 26 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 27 Select **HDFS** in the required cluster from the **Service**.

Step 28 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 29 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.106 ALM-14012 JournalNode Is Out of Synchronization

Description

On the active NameNode, the system checks the data consistency of all JournalNodes in the cluster every 5 minutes. This alarm is generated when the data on a JournalNode is inconsistent with the data on the other JournalNodes.

This alarm is cleared in 5 minutes after the data on JournalNodes is consistent.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14012	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NameServiceName	Specifies the NameService for which the alarm is generated.

Impact on the System

When a JournalNode is working incorrectly, the data on the node becomes inconsistent with that on the other JournalNodes. If data on more than half of JournalNodes is inconsistent, the NameNode cannot work correctly, making the HDFS service unavailable.

Possible Causes

- The JournalNode instance does not exist (deleted or migrated).
- The JournalNode instance has not been started or has been stopped.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The JournalNode instance is working incorrectly.
- The network of the JournalNode is unreachable.

Procedure

Check whether the JournalNode instance has been started up.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. In the alarm list, click the alarm.

Step 2 Check **Location** and obtain the IP address of the JournalNode for which the alarm is generated.

Step 3 Choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Instance**. In the instance list, check whether the JournalNode instance exists on the node for which the alarm is generated.

- If yes, go to **Step 5**.
- If no, go to **Step 4**.

Step 4 Choose **O&M** > **Alarm** > **Alarms**. In the alarm list, click **Clear** in the **Operation** column of the alarm. In the dialog box that is displayed, click **OK**. No further action is needed.

Step 5 Click the JournalNode instance and check whether its **Configuration Status** is **Synchronized**.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Select the JournalNode instance and choose **Start Instance** to start the instance.

Step 7 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Check whether the JournalNode instance is working correctly.

Step 8 Check whether **Running Status** of the JournalNode instance is **Normal**.

- If yes, go to **Step 11**.
- If no, go to **Step 9**.

Step 9 Select the JournalNode instance and choose **More** > **Restart Instance** to start the instance.

NOTE

If the number of JournalNode instances restarted exceeds one third of the total number of JournalNodes, the HDFS service may be faulty.

Step 10 After 5 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Check whether the network of the JournalNode is reachable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Instance** to check the service IP address of the active NameNode.

Step 12 Log in to the active NameNode as user **root**.

Step 13 Run the **ping** command to check whether a timeout occurs or the network is unreachable between the active NameNode and the JournalNode.

ping *service IP address of the JournalNode*

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Contact the network administrator to rectify the network fault and check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 15**.

Collect fault information.

Step 15 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 16 Select **HDFS** in the required cluster from the **Service**.

Step 17 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 18 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.107 ALM-14013 Failed to Update the NameNode FsImage File

Alarm Description

HDFS metadata is stored in the FsImage file of the NameNode data directory, which is specified by the **dfs.namenode.name.dir** configuration item. The standby NameNode periodically combines existing FsImage files and Editlog files stored in the JournalNode to generate a new FsImage file, and then pushes the new FsImage file to the data directory of the active NameNode. This period is specified by the **dfs.namenode.checkpoint.period** configuration item of HDFS. The default value is 3600s, namely, one hour. If the FsImage file in the data directory of the active NameNode is not updated, the HDFS metadata combination function is abnormal and requires rectification.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

On the active NameNode, the system checks the FsImage file information every five minutes. This alarm is generated when no FsImage file is generated within three combination periods.

This alarm is cleared when a new FsImage file is generated and pushed to the active NameNode, which indicates that the HDFS metadata combination function can be properly used.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14013	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NameServiceName	Specifies the NameService for which the alarm is generated.

Impact on the System

If the FsImage file in the data directory of the active NameNode is not updated, the HDFS metadata combination function is abnormal and requires rectification. If it is not rectified, the Editlog files increase continuously after HDFS runs for a period. In this case, HDFS restart is time-consuming because a large number of Editlog files need to be loaded. In addition, this alarm also indicates that the standby NameNode is abnormal and the NameNode high availability (HA) mechanism becomes invalid. When the active NameNode is faulty, the HDFS service becomes unavailable.

Possible Causes

- The standby NameNode is stopped.
- The standby NameNode instance is working incorrectly.
- The standby NameNode fails to generate a new FsImage file.
- Space of the data directory on the standby NameNode is insufficient.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The standby NameNode fails to push the FsImage file to the active NameNode.
- Space of the data directory on the active NameNode is insufficient.

Handling Procedure

Check whether the standby NameNode is stopped.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. In the alarm list, click the alarm.

Step 2 View **Location** and obtain the host name of the active NameNode for which the alarm is generated and name of the NameService where the active NameNode resides.

Step 3 Choose **Cluster** > **Services** > **HDFS** > **Instances**, find the standby NameNode instance of the NameService in the instance list, and check whether its **Configuration Status** is **Synchronized**.

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Select the standby NameNode instance, choose **Start Instance**, and wait until the startup is complete.

Step 5 Wait for a NameNode metadata combination period and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the NameNode instance is working correctly.

Step 6 Check whether **Running Status** of the standby NameNode instance is **Normal**.

- If yes, go to **Step 9**.
- If no, go to **Step 7**.

Step 7 Select the standby NameNode instance, choose **More** > **Restart Instance**, and wait until the startup is complete.

NOTE

Services are not affected after the standby NameNode is restarted.

Step 8 Wait for a NameNode metadata combination period and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 30**.

Check whether the standby NameNode fails to generate a new FsImage file.

Step 9 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS** > **Configurations** > **All Configurations**, and search and obtain the value of **dfs.namenode.checkpoint.period**. This value is the period of NameNode metadata.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Choose **Cluster** > **Services** > **HDFS** > **Instance** and obtain the service IP addresses of the active and standby NameNodes of the NameService for which the alarm is generated.

Step 11 Click the **NameNode(*xx*,Standby)** and **Instance Configurations** to obtain the value of **dfs.namenode.name.dir**. This value is the FsImage storage directory of the standby NameNode.

Step 12 Log in to the standby NameNode node as user **root**.

Run the following command to switch to user **omm**:

```
su - omm
```

Step 13 Go to the FsImage storage directory and check the generation time of the newest FsImage file.

```
cd Storage directory of the standby NameNode/current
stat -c %y $(ls -t | grep "fsimage_[0-9]*$" | head -1)
```

Step 14 Run the following command to obtain the current system time:

```
date
```

Step 15 Calculate the time difference between the generation time of the newest FsImage file and the current system time and check whether the time difference is greater than three times of the metadata combination period (**dfs.namenode.checkpoint.period**).

- If yes, go to **Step 16**.
- If no, go to **Step 20**.

Step 16 The metadata combination function of the standby NameNode is faulty. Run the following command to check whether the fault is caused by insufficient storage space.

Go to the FsImage storage directory.

```
cd Storage directory of the standby NameNode/current
```

Check the size of the newest FsImage file (in MB).

```
du -m $(ls -t | grep "fsimage_[0-9]*$" | head -1) | awk '{print $1}'
```

Step 17 Run the following command to check the available disk space of the standby NameNode (in MB).

```
df -m ./ | awk 'END{print $4}'
```

Step 18 Compare the FsImage file size and the available disk space and determine whether another FsImage file can be stored on the disk.

- If yes, go to **Step 7**.
- If no, go to **Step 19**.

Step 19 Clear the redundant files on the disk where the directory resides to reserve sufficient space for metadata. After the clearance, wait for a NameNode metadata combination period and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 20**.

Check whether the standby NameNode fails to push the FsImage file to the active NameNode.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 20 Log in to the standby NameNode node as user **root**.

Step 21 Run the following command to switch to user **omm**:

```
su - omm
```

Step 22 Run the following command to check whether the standby NameNode can push the file to the active NameNode.

```
tmpFile=/tmp/tmp_test_$(date +%s)
echo "test" > $tmpFile
scp $tmpFile IP address of the active NameNode node:/tmp
```

- If yes, go to **Step 24**.
- If no, go to **Step 23**.

Step 23 When the standby NameNode fails to push data to the active NameNode as user **omm**, contact the MRS system administrator to handle the fault. Wait for a NameNode metadata combination period and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 24**.

Check whether space on the data directory of the active NameNode is insufficient.

Step 24 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Instances**, click the active NameNode of the NameService for which the alarm is generated, and then click **Instance Configurations** to obtain the value of **dfs.namenode.name.dir**. This value is the FsImage storage directory of the active NameNode.

Step 25 Log in to the active NameNode node as user **root** or **omm**:

Step 26 Go to the FsImage storage directory and check the size of the newest FsImage file (in MB).

```
cd Storage directory of the active NameNode node/current
du -m $(ls -t | grep "fsimage_[0-9]*$" | head -1) | awk '{print $1}'
```

Step 27 Run the following command to check the available disk space of the active NameNode (in MB).

```
df -m ./ | awk 'END{print $4}'
```

Step 28 Compare the FsImage file size and the available disk space and determine whether another FsImage file can be stored on the disk.

- If yes, go to **Step 30**.
- If no, go to **Step 29**.

Step 29 Clear the redundant files on the disk where the directory resides to reserve sufficient space for metadata. After the clearance, wait for a NameNode metadata combination period and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 30**.

Collect fault information.

Step 30 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 31 Select **NameNode** in the required cluster from the **Service**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 32 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 33 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.108 ALM-14014 NameNode GC Time Exceeds the Threshold

Description

The system checks the garbage collection (GC) duration of the NameNode process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default).

This alarm is cleared when the GC duration is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14014	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A long GC duration of the NameNode process may interrupt the services and users cannot read or write files.

Possible Causes

The heap memory of the NameNode instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Procedure

Check the GC duration.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms**. On the displayed interface, click the drop-down button of **ALM-14014 NameNode GC Time Exceeds the Threshold.** Then check the role name in **Location** and confirm the IP address of the instance.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **HDFS**. On the page that is displayed, click the **Instance** tab. In the instance list, select **NameNode** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Garbage Collection**, and select **NameNode Garbage Collection (GC)**. Check the GC duration statistics of the NameNode process every minute.

Step 3 Check whether the GC duration of the NameNode process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **HDFS** > **Configurations** > **All Configurations** > **NameNode** > **System**. Increase the value of the **GC_OPTS** parameter as required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The recommended mapping between the number of HDFS file objects (filesystem objects = files + blocks) and the JVM parameters configured for NameNode is as follows:

- If the number of file objects reaches 10,000,000, you are advised to set the JVM parameters as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
- If the number of file objects reaches 20,000,000, you are advised to set the JVM parameters as follows: -Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G
- If the number of file objects reaches 50,000,000, you are advised to set the JVM parameters as follows: -Xms32G -Xmx32G -XX:NewSize=3G -XX:MaxNewSize=3G
- If the number of file objects reaches 100,000,000, you are advised to set the JVM parameters as follows: -Xms64G -Xmx64G -XX:NewSize=6G -XX:MaxNewSize=6G
- If the number of file objects reaches 200,000,000, you are advised to set the JVM parameters as follows: -Xms96G -Xmx96G -XX:NewSize=9G -XX:MaxNewSize=9G
- If the number of file objects reaches 300,000,000, you are advised to set the JVM parameters as follows: -Xms164G -Xmx164G -XX:NewSize=12G -XX:MaxNewSize=12G

Step 5 Save the configuration and restart the NameNode instance.

NOTE

- During the restart of the active NameNode, a NameNode active/standby switchover occurs. As a result, no active node is available in the system for a short period of time (in the transition period of the active/standby switchover), and an alarm indicating that the HDFS service is unavailable may be generated. In addition, an error is reported during the running read and write tasks, but services are not interrupted.
- Services are not affected after the standby NameNode is restarted.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **NameNode** in the required cluster from the **Service**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.109 ALM-14015 DataNode GC Time Exceeds the Threshold

Description

The system checks the garbage collection (GC) duration of the DataNode process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default).

This alarm is cleared when the GC duration is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14015	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A long GC duration of the DataNode process may interrupt the services and users cannot read or write files.

Possible Causes

The heap memory of the DataNode instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the GC duration.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms**. On the displayed interface, click the drop-down button of **ALM-14015 DataNode GC Time Exceeds the Threshold**. Then check the role name in **Location** and confirm the IP address of the instance.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **HDFS**. On the page that is displayed, click the **Instance** tab. In the instance list, select **DataNode** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Garbage Collection**, and select **DataNode Garbage Collection (GC)**. Check the GC duration statistics of the DataNode process every minute.

Step 3 Check whether the GC duration of the DataNode process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **HDFS** > **Configurations** > **All Configurations** > **DataNode** > **System**. Increase the value of the **GC_OPTS** parameter as required.

NOTE

The mapping between the average number of blocks of a DataNode instance and the DataNode memory is as follows:

- If the average number of blocks of a DataNode instance reaches 2,000,000, the reference values of the JVM parameters of the DataNode are as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
- If the average number of blocks of a DataNode instance reaches 5,000,000, the reference values of the JVM parameters of the DataNode are as follows: -Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G

Step 5 Save the configuration and restart the DataNode instance.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform this operation during off-peak hours.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **DataNode** in the required cluster from the **Service**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.110 ALM-14016 DataNode Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of HDFS every 30 seconds. This alarm is generated when the direct memory usage of DataNode instances exceeds the threshold (90% of the maximum memory).

This alarm is automatically cleared when the direct memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14016	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the available direct memory of DataNode instances is insufficient, a memory overflow may occur and the service breaks down.

Possible Causes

The direct memory of DataNode instances is overused or the direct memory is inappropriately allocated.

Handling Procedure

Check the direct memory usage.

Step 1 On the **Home** page of FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, click the drop-down list in the row containing **ALM-14016 DataNode Direct Memory Usage Exceeds the Threshold**, and view the role name and IP address of the instance for which the alarm is generated in the **Location** area.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > **Services** > **HDFS**. On the page that is displayed, click the **Instance** tab. In the instance list, select **DataNode** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart, choose **Customize** > **Resource**, and select **DataNode Memory** to check the direct memory usage.

Step 3 Check whether the used direct memory of a DataNode instance reaches 90% (default threshold) of the maximum direct memory allocated to it.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **HDFS** > **Configurations** > **All Configurations** > **DataNode** > **System**. Check whether **-XX:MaxDirectMemorySize** exists in the **GC_OPTS** parameter.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Adjust the value of **-XX:MaxDirectMemorySize**.

1. In **GC_OPTS**, check the value of **-Xmx** and check whether the node memory is sufficient.

   NOTE

   You can determine whether the node memory is sufficient based on the actual environment. For example, you can use the following method:

   Use the IP address to log in to the instance for which the alarm is generated as user **root** and run the **free -g** command to check the value of **Mem** in the **free** column. The value indicates the available memory of the node. In the following example, the available memory of the node is 4 GB.

   ```
                total        used        free      shared  buff/cache   available
   Mem:            112          48           4          10          58          46
   ......
   ```

   If the value of **Mem** is at least that of **-Xmx**, the node memory is sufficient. If the value of **Mem** is less than that of **-Xmx**, the node memory is insufficient.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, change the value of **-XX:MaxDirectMemorySize** to that of **-Xmx**.
- If no, increase **-XX:MaxDirectMemorySize** to a value no larger than that of **Mem**.

2. Save the configuration and restart the DataNode instances.

   NOTE

   Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

Step 6 Check whether **ALM-14008 DataNode Heap Memory Usage Exceeds the Threshold** exists.

- If yes, rectify the fault by referring to **ALM-14008 DataNode Heap Memory Usage Exceeds the Threshold**.
- If no, go to **Step 7**.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **DataNode** for the target cluster.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.111 ALM-14017 NameNode Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the HDFS service every 30 seconds. This alarm is generated when the direct memory usage of a NameNode instance exceeds the threshold (90% of the maximum memory).

The alarm is cleared when the direct memory usage is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14017	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available direct memory of the HDFS service is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The direct memory of the NameNode instance is overused or the direct memory is inappropriately allocated.

Procedure

Check the direct memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms.** On the displayed interface, click the drop-down button of **ALM-14017 NameNode Direct Memory Usage Exceeds the Threshold**. Then check the role name in **Location** and confirm the IP address of the instance.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **HDFS** and click the **Instance** tab. On the displayed page, select **NameNode** (the IP address of the instance for which the alarm is generated). Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **Resource**, select **NameNode Memory**. Check the direct memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the used direct memory of NameNode reaches 90% of the maximum direct memory specified for NameNode by default.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Configurations** > **NameNode** > **All Configurations**. Check whether **-XX:MaxDirectMemorySize** exists in the **GC_OPTS** parameter.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 In the **GC_OPTS** parameter, delete "-XX:MaxDirectMemorySize". Save the configuration and restart the NameNode instance.

NOTE

- During the restart of the active NameNode, a NameNode active/standby switchover occurs. As a result, no active node is available in the system for a short period of time (in the transition period of the active/standby switchover), and an alarm indicating that the HDFS service is unavailable may be generated. In addition, an error is reported during the running read and write tasks, but services are not interrupted.
- Services are not affected after the standby NameNode is restarted.

Step 6 Check whether the **ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold** alarm exists.

- If yes, handle the alarm by referring to **ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold**.
- If no, go to **Step 7**.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 9 Select **NameNode** in the required cluster from the **Service**.

Step 10 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.112 ALM-14018 NameNode Non-heap Memory Usage Exceeds the Threshold

Description

The system checks the non-heap memory usage of the HDFS NameNode every 30 seconds and compares the actual usage with the threshold. The non-heap memory usage of the HDFS NameNode has a default threshold. This alarm is generated when the non-heap memory usage of the HDFS NameNode exceeds the threshold.

Users can choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* > **HDFS** to change the threshold.

This alarm is cleared when the no-heap memory usage of the HDFS NameNode is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14018	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the memory usage of the HDFS NameNode is too high, data read/write performance of HDFS will be affected.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

Non-heap memory of the HDFS NameNode is insufficient.

Procedure

Delete unnecessary files.

Step 1 Log in to the HDFS client as user **root**. Run the **cd** command to go to the client installation directory, and run the **source bigdata_env** command.

If the cluster adopts the security mode, perform security authentication.

Run the **kinit hdfs** command and enter the password as prompted. Obtain the password from the administrator.

Step 2 Run the **hdfs dfs -rm -r** *file or directory path* command to delete unnecessary files.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the NameNode JVM non-heap memory usage and configuration.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS**. The HDFS status page is displayed.

Step 5 In the **Basic Information** area, click **NameNode(Active)**. The HDFS WebUI is displayed.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 6 On the HDFS WebUI, click the **Overview** tab. In **Summary**, check the numbers of files, directories, and blocks in HDFS.

Step 7 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Configurations** > **All Configurations**. In **Search**, enter **GC_OPTS** to check the **GC_OPTS** non-heap memory parameter of **HDFS->NameNode**.

Adjust system configurations.

Step 8 Check whether the non-heap memory is properly configured based on the number of file objects in **Step 6** and the non-heap parameters configured for NameNode in **Step 7**.

- If yes, go to **Step 9**.
- If no, go to **Step 12**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The recommended mapping between the number of HDFS file objects (filesystem objects = files + blocks) and the JVM parameters configured for NameNode is as follows:

- If the number of file objects reaches 10,000,000, you are advised to set the JVM parameters as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
- If the number of file objects reaches 20,000,000, you are advised to set the JVM parameters as follows: -Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G
- If the number of file objects reaches 50,000,000, you are advised to set the JVM parameters as follows: -Xms32G -Xmx32G -XX:NewSize=3G -XX:MaxNewSize=3G
- If the number of file objects reaches 100,000,000, you are advised to set the JVM parameters as follows: -Xms64G -Xmx64G -XX:NewSize=6G -XX:MaxNewSize=6G
- If the number of file objects reaches 200,000,000, you are advised to set the JVM parameters as follows: -Xms96G -Xmx96G -XX:NewSize=9G -XX:MaxNewSize=9G
- If the number of file objects reaches 300,000,000, you are advised to set the JVM parameters as follows: -Xms164G -Xmx164G -XX:NewSize=12G -XX:MaxNewSize=12G

Step 9 Modify the **GC_OPTS** parameter of the NameNode based on the mapping between the number of file objects and non-heap memory.

Step 10 Save the configuration and click **Dashboard** > **More** > **Restart Service**.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 11 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 13 Select the following services in the required cluster from the **Service**.

- ZooKeeper
- HDFS

Step 14 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 15 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.113 ALM-14019 DataNode Non-heap Memory Usage Exceeds the Threshold

Description

The system checks the non-heap memory usage of the HDFS DataNode every 30 seconds and compares the actual usage with the threshold. The non-heap memory usage of the HDFS DataNode has a default threshold. This alarm is generated when the non-heap memory usage of the HDFS DataNode exceeds the threshold.

Users can choose **O&M > Alarm > Thresholds>** *Name of the desired cluster* **> HDFS** to change the threshold.

This alarm is cleared when the no-heap memory usage of the HDFS DataNode is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14019	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the memory usage of the HDFS DataNode is too high, data read/write performance of HDFS will be affected.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

Non-heap memory of the HDFS DataNode is insufficient.

Procedure

Delete unnecessary files.

Step 1 Log in to the HDFS client as user **root**. Run the **cd** command to go to the client installation directory, and run the **source bigdata_env** command.

If the cluster adopts the security mode, perform security authentication.

Run the **kinit hdfs** command and enter the password as prompted. Obtain the password from the administrator.

Step 2 Run the **hdfs dfs -rm -r** *file or directory path* command to delete unnecessary files.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the DataNode JVM non-heap memory usage and configuration.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS**.

Step 5 In the **Basic Information** area, click **NameNode(Active)**. The HDFS WebUI is displayed.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 6 On the HDFS WebUI, click the **Datanodes** tab to view the number of blocks of all DataNodes that report alarms.

Step 7 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Configurations** > **All Configurations**. In **Search**, enter **GC_OPTS** to check the **GC_OPTS** non-heap memory parameter of **HDFS->DataNode**.

Adjust system configurations.

Step 8 Check whether the memory is properly configured based on the number of blocks in **Step 6** and the memory parameters configured for DataNode in **Step 7**.

- If yes, go to **Step 9**.
- If no, go to **Step 12**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The mapping between the average number of blocks of a DataNode instance and the DataNode memory is as follows:

- If the average number of blocks of a DataNode instance reaches 2,000,000, the reference values of the JVM parameters of the DataNode are as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
- If the average number of blocks of a DataNode instance reaches 5,000,000, the reference values of the JVM parameters of the DataNode are as follows: -Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G

Step 9 Modify the **GC_OPTS** parameter of the DataNode based on the mapping between the number of blocks and memory.

Step 10 Save the configuration and click **Dashboard** > **More** > **Restart Service**.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 11 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 13 Select the following services in the required cluster from the **Service**.

- ZooKeeper
- HDFS

Step 14 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 15 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.114 ALM-14020 Number of Entries in the HDFS Directory Exceeds the Threshold

Description

The system obtains the number of subfiles and subdirectories in a specified directory every hour and checks whether it reaches the percentage of the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

threshold (the maximum number of subfiles and subdirectories in an HDFS directory, the threshold for triggering an alarm is **90%** by default). If it exceeds the percentage of the threshold, an alarm is triggered.

When the number of subfiles and subdirectories in the directory the alarm is lower than the percentage of the threshold, the alarm is automatically cleared. When the monitoring switch is disabled, alarms corresponding to all directories are cleared. If a directory is removed from the monitoring list, alarms corresponding to the directory are cleared.

NOTE

- The **dfs.namenode.fs-limits.max-directory-items** parameter specifies the maximum number of subfiles and subdirectories in the HDFS directory. Its default value is **1048576**. If the number of subfiles and subdirectories in a directory exceeds the parameter value, subfiles and subdirectories cannot be created in the directory.
- The **dfs.namenode.directory-items.monitor** parameter specifies the list of directories to be monitored. Its default value is **/tmp,/SparkJobHistory,/mr-history**.
- The **dfs.namenode.directory-items.monitor.enabled** parameter is used to enable or disable the monitoring switch. Its default value is **true**, which means the monitoring switch is enabled by default.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14020	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
NameServiceName	Specifies the NameService service for which the alarm is generated.
Directory	Specifies the directory for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the number of entries in the monitored directory exceeds 90% of the threshold, an alarm is triggered, but entries can be added to the directory. Once the maximum threshold is exceeded, entries will fail to be added to the directory.

Possible Causes

The number of entries in the monitored directory exceeds 90% of the threshold.

Procedure

Check whether unnecessary files exist in the system.

Step 1 Log in to the HDFS client as user **root**. Run **cd** to switch to the client installation directory, and run **source bigdata_env** to configure the environment variables.

If the cluster is in security mode, security authentication is required.

Run the **kinit hdfs** command and enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 2 Run the following command to check whether files and directories in the directory with the alarm can be deleted:

hdfs dfs -ls *Directory with the alarm*

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Run the following command to delete unnecessary files.

hdfs dfs -rm -r -f *File or directory path*

NOTE

Deleting a file or folder is a high-risk operation. Ensure that the file or folder is no longer required before performing this operation.

Step 4 Wait 1 hour and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the threshold is correctly configured.

Step 5 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **HDFS** > **Configurations** > **All Configurations**. Search for the **dfs.namenode.fs-limits.max-directory-items** parameter and check whether the parameter value is appropriate.

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 Increase the parameter value.

Step 7 Save the configuration and click **Dashboard** > **More** > **Restart Service**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 8 Wait 1 hour and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 10 Select **HDFS** in the required cluster from the **Service**.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.115 ALM-14021 NameNode Average RPC Processing Time Exceeds the Threshold

Alarm Description

The system checks the average RPC processing time of NameNode every 30 seconds, and compares the actual average RPC processing time with the threshold (default value: 100 ms). This alarm is generated when the system detects that the average RPC processing time exceeds the threshold for several consecutive times (10 times by default).

You can choose **O&M** > **Alarm** > **Thresholds** > **HDFS** and change the threshold.

When **Trigger Count** is **1**, this alarm is cleared when the average RPC processing time of NameNode is less than or equal to the threshold. When **Trigger Count** is greater than **1**, this alarm is cleared when the average RPC processing time of NameNode is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14021	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NameServiceName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

NameNode cannot process the RPC requests from HDFS clients, upper-layer services that depend on HDFS, and DataNode in a timely manner. Specifically, the services that access HDFS run slowly or the HDFS service is unavailable.

Possible Causes

- The CPU performance of NameNode nodes is insufficient and therefore NameNode nodes cannot process messages in a timely manner.
- The configured NameNode memory is too small and frame freezing occurs on the JVM due to frequent full garbage collection.
- NameNode parameters are not configured properly, so NameNode cannot make full use of system performance.

Handling Procedure

Obtain alarm information.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. In the alarm list, click the alarm.

Step 2 Check the alarm. Obtain the host name of the NameNode node involved in this alarm from the **HostName** information of **Location**. Then obtain the name of the NameService node involved in this alarm from the **NameServiceName** information of **Location**.

Check whether the threshold is too small.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check the status of the services that depend on HDFS. Check whether the services run slowly or task execution times out.

- If yes, go to **Step 8**.
- If no, go to **Step 4**.

Step 4 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS**. Click the drop-down menu in the upper right corner of the **Chart** area, choose **Customize** > **RPC**, select **Average Time of Active NameNode RPC Processing**, and click **OK**.

Step 5 On the **Average Time of Active NameNode RPC Processing** monitoring page, obtain the value of the NameService node involved in this alarm.

Step 6 On the FusionInsight Manager homepage, choose **O&M** > **Alarm** > **Thresholds** > **HDFS**. Locate **Average Time of Active NameNode RPC Processing** and click the **Modify** in the **Operation** column of the default rule. The **Modify Rule** page is displayed. Change **Threshold** to 150% of the peak value within one day before and after the alarm is generated. Click **OK** to save the new threshold.

Step 7 Wait for 5 minutes and then check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the CPU performance of the NameNode node is sufficient.

Step 8 On the FusionInsight Manager portal, click **O&M > Alarm >Alarms** and check whether **ALM-12016 CPU Usage Exceeds the Threshold** is generated for the NameNode node whose ID is 12016.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Rectify the fault by following the handling procedure of **ALM-12016 CPU Usage Exceeds the Threshold**.

Step 10 Wait for 10 minutes and check whether alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check whether the memory of the NameNode node is too small.

Step 11 On the FusionInsight Manager portal, click **O&M > Alarm >Alarms** and check whether **ALM-14007 HDFS NameNode Heap Memory Usage Exceeds the Threshold** is generated for the NameNode node whose ID is 14007.

- If yes, go to **Step 12**.
- If no, go to **Step 14**.

Step 12 Rectify the fault by following the handling procedure of **ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold**.

Step 13 Wait for 10 minutes and check whether alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Check whether NameNode parameters are configured properly.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 14 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Configurations** > **All Configurations**. Search for parameter **dfs.namenode.handler.count** and view its value. If the value is less than or equal to 128, change it to **128**. If the value is greater than 128 but less than 192, change it to **192**.

Step 15 Search for parameter **ipc.server.read.threadpool.size** and view its value. If the value is less than 5, change it to **5**.

Step 16 Click **Save** and click **OK**.

Step 17 On the **Instance** page of HDFS, select the standby NameNode of NameService involved in this alarm and choose **More** > **Restart Instance**. Enter the password and click **OK**. Wait until the standby NameNode is started up.

NOTE

Services are not affected after the standby NameNode is restarted.

Step 18 On the **Instance** page of HDFS, select the active NameNode of NameService involved in this alarm and choose **More** > **Restart Instance**. Enter the password and click **OK**. Wait until the active NameNode is started up.

Step 19 Wait for 1 hour and then check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 20**.

Collect fault information.

Step 20 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 21 Select the following node in the required cluster from the **Service**.

- HDFS

Step 22 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 23 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.116 ALM-14022 NameNode Average RPC Queuing Time Exceeds the Threshold

Alarm Description

The system checks the average RPC queuing time of NameNode every 30 seconds, and compares the actual average RPC queuing time with the threshold (default

value: 200 ms). This alarm is generated when the system detects that the average RPC queuing time exceeds the threshold for several consecutive times (10 times by default).

You can choose **O&M** > **Alarm** > **Thresholds** > **HDFS** and change the threshold.

When **Trigger Count** is **1**, this alarm is cleared when the average RPC queuing time of NameNode is less than or equal to the threshold. When **Trigger Count** is greater than **1**, this alarm is cleared when the average RPC queuing time of NameNode is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14022	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NameServiceName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

NameNode cannot process the RPC requests from HDFS clients, upper-layer services that depend on HDFS, and DataNode in a timely manner. Specifically, the services that access HDFS run slowly or the HDFS service is unavailable.

Possible Causes

- The CPU performance of NameNode nodes is insufficient and therefore NameNode nodes cannot process messages in a timely manner.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The configured NameNode memory is too small and frame freezing occurs on the JVM due to frequent full garbage collection.
- NameNode parameters are not configured properly, so NameNode cannot make full use of system performance.
- The volume of services that access HDFS is too large and therefore NameNode is overloaded.

Handling Procedure

Obtain alarm information.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. In the alarm list, click the alarm.

Step 2 Check the alarm. Obtain the alarm generation time from **Generated**. Obtain the host name of the NameNode node involved in this alarm from the **HostName** information of **Location**. Then obtain the name of the NameService node involved in this alarm from the **NameServiceName** information of **Location**.

Check whether the threshold is too small.

Step 3 Check the status of the services that depend on HDFS. Check whether the services run slowly or task execution times out.

- If yes, go to **Step 8**.
- If no, go to **Step 4**.

Step 4 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS**. Click the drop-down menu in the upper right corner of **Chart**, choose **Customize**, select **Average Time of Active NameNode RPC Queuing**, and click **OK**.

Step 5 On the **Average Time of Active NameNode RPC Queuing** monitoring page, obtain the value of the NameService node involved in this alarm.

Step 6 On the FusionInsight Manager homepage, choose **O&M** > **Alarm** > **Thresholds** > **HDFS**. Locate **Average Time of Active NameNode RPC Queuing** and click the **Modify** in the **Operation** column of the default rule. The **Modify Rule** page is displayed. Change **Threshold** to 150% of the monitored value. Click **OK** to save the new threshold.

Step 7 Wait for 1 minute and then check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the CPU performance of the NameNode node is sufficient.

Step 8 On the FusionInsight Manager portal, click **O&M > Alarm > Alarms** and check whether **ALM-12016 CPU Usage Exceeds the Threshold** is generated.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Rectify the fault by following the handling procedure of **ALM-12016 CPU Usage Exceeds the Threshold**.

Step 10 Wait for 10 minutes and check whether alarm 14022 is automatically cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check whether the memory of the NameNode node is too small.

Step 11 On the FusionInsight Manager portal, click **O&M > Alarm > Alarms** and check whether **ALM-14007 HDFS NameNode Memory Usage Exceeds the Threshold** is generated.

- If yes, go to **Step 12**.
- If no, go to **Step 14**.

Step 12 Rectify the fault by following the handling procedure of **ALM-14007 NameNode Heap Memory Usage Exceeds the Threshold**.

Step 13 Wait for 10 minutes and check whether alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Check whether NameNode parameters are configured properly.

Step 14 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS** > **Configurations** > **All Configurations**. Search for parameter **dfs.namenode.handler.count** and view its value.

If the value is lower than or equal to 128, change it to **128**. If the value is greater than 128 but less than 192, change it to **192**.

Step 15 Search for parameter **ipc.server.read.threadpool.size** and view its value.

If the value is less than 5, change it to **5**.

Step 16 Click **Save** and then **OK** to save the settings.

Step 17 On the **Instance** page of HDFS, select the standby NameNode of NameService involved in this alarm and choose **More** > **Restart Instance**. Enter the required password and click **OK** . Wait until the standby NameNode is started up.

NOTE

Services are not affected after the standby NameNode is restarted.

Step 18 On the **Instance** page of HDFS, select the active NameNode of NameService involved in this alarm and choose **More** > **Restart Instance**. Enter the required password and click **OK**. Wait until the active NameNode is started up.

Step 19 Wait for 1 hour and then check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 20**.

Check whether the HDFS workload changes and reduce the workload properly.

Step 20 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HDFS**. Click the drop-down menu in the upper right corner of the **Chart** area, choose **Customize** > **RPC**, select **Average Time of Active NameNode RPC Queuing**, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 21 Click . The **Details** page is displayed.

Step 22 Set the monitoring data display period, from 5 days before the alarm generation time to the alarm generation time. Click **OK**.

Step 23 On the **Average RPC Queuing Time** monitoring page, check whether the point in time when the queuing time increases abruptly exists.

- If yes, go to **Step 24**.
- If no, go to **Step 27**.

Step 24 Confirm and check the point in time. Check whether a new task frequently accesses HDFS and whether the access frequency can be reduced.

Step 25 If a Balancer task starts at the point in time, stop the task or specify a node for the task to reduce the HDFS workload.

Step 26 Wait for 1 hour and then check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 27**.

Collect fault information.

Step 27 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 28 Select **HDFS** in the required cluster from the **Service**.

Step 29 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 30 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.117 ALM-14023 Percentage of Total Reserved Disk Space for Replicas Exceeds the Threshold

Description

The system checks the percentage of total reserved disk space for replicas (Total reserved disk space for replicas/(Total reserved disk space for replicas + Total remaining disk space)) every 30 seconds and compares the actual percentage with the threshold (**90%** by default). This alarm is generated when the percentage of total reserved disk space for replicas exceeds the threshold for multiple consecutive times (**Trigger Count**).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The alarm is cleared in the following two scenarios: The value of **Trigger Count** is **1** and the percentage of total reserved disk space for replicas is less than or equal to the threshold; the value of **Trigger Count** is greater than **1** and the percentage of total reserved disk space for replicas is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14023	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
NameServiceName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The performance of writing data to HDFS is affected. If all remaining DataNode space is reserved for replicas, writing HDFS data fails.

Possible Causes

- The alarm threshold is improperly configured.
- The disk space configured for the HDFS cluster is insufficient.
- The volume of services that access HDFS is too large and therefore DataNode is overloaded.

Procedure

Check whether the alarm threshold is appropriate.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* > **HDFS** > **Disk** > **Percentage of Reserved Space for**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Replicas of Unused Space to check whether the alarm threshold is appropriate. (The default threshold is **90%**. Users can change it as required.)

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* > **HDFS** > **Disk** > **Percentage of Reserved Space for Replicas of Unused Space** and Click **Modify,** change the threshold based on the actual usage.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether an alarm indicating insufficient disk space is generated.

Step 4 On the FusionInsight Manager portal, check whether **ALM-14001 HDFS Disk Usage Exceeds the Threshold** or **ALM-14002 DataNode Disk Usage Exceeds the Threshold** exists on the **O&M > Alarm > Alarms** page.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Rectify the fault by referring to **ALM-14001 HDFS Disk Usage Exceeds the Threshold** or **ALM-14002 DataNode Disk Usage Exceeds the Threshold**. Check whether the alarm is cleared.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Expand the DataNode capacity.

Step 7 Expand the DataNode capacity .

Step 8 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 10 Select **HDFS** in the required cluster from the **Service**.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 20 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.118 ALM-14024 Tenant Space Usage Exceeds the Threshold

Description

The system checks the space usage (used space of each directory/space allocated to each directory) of each directory associated with a tenant every hour and compares the space usage of each directory with the threshold set for the directory. This alarm is generated when the space usage exceeds the threshold.

This alarm is cleared when the space usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14024	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
TenantName	Specifies the tenant for which the alarm is generated.
DirectoryName	Specifies the directory for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

This alarm is generated if the space usage of the tenant directory exceeds the custom threshold. File writing to the directory is not affected. If the used space exceeds the maximum storage space allocated to the directory, the HDFS fails to write data to the directory.

Possible Causes

- The alarm threshold is improperly configured.
- The space allocated to the tenant is improper.

Procedure

Check whether the alarm threshold is appropriate.

Step 1 View the alarm location information to obtain the tenant name and tenant directory for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose the **Tenant Resources** page, select the tenant for which the alarm is generated, and click **Resources**. Check whether the storage space threshold configured for the tenant directory for which the alarm is generated is proper. (The default value 90% is a proper value. You can set it based on the site requirements.)

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 On the **Resources** page, click **Modify** to modify or delete the storage space threshold.

Step 4 About one minute later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the space allocated to the tenant is appropriate.

Step 5 On the FusionInsight Manager portal, choose the **Tenant Resources** page, select the tenant for which the alarm is generated, and click **Resources**. Check whether the storage space quota of the tenant directory for which the alarm is generated is proper based on the actual service status of the tenant directory.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 On the **Resources** page, click **Modify** to modify the storage space quota.

Step 7 About one minute later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 9 Select **HDFS** in the required cluster and **NodeAgent** under **Manager** from the **Service**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 20 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.119 ALM-14025 Tenant File Object Usage Exceeds the Threshold

Description

The system checks the file object usage (used file objects of each directory/number of file objects allocated to each directory) of each directory associated with a tenant every hour and compares the file object usage of each directory with the threshold set for the directory. This alarm is generated when the file object usage exceeds the threshold.

This alarm is cleared when the file object usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
14025	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
TenantName	Specifies the tenant for which the alarm is generated.
DirectoryName	Specifies the directory for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm.

Impact on the System

This alarm is generated if the usage of file objects in a tenant directory exceeds the custom threshold. File writing to the directory is not affected. If the number of used file objects exceeds the maximum number of file objects allocated to the directory, the HDFS fails to write data to the directory.

Possible Causes

- The alarm threshold is improperly configured.
- The maximum number of file objects allocated to the tenant directory is inappropriate.

Procedure

Check whether the alarm threshold is appropriate.

Step 1 View the alarm location information to obtain the tenant name and tenant directory for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose the **Tenant Resources** page, select the tenant for which the alarm is generated, and click **Resources**. Check whether the file object threshold configured for the tenant directory for which the alarm is generated is proper. (The default value 90% is a proper value. You can set it based on the site requirements.)

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 On the **Resources** page, click **Modify** to modify or delete the file object threshold of the tenant directory for which the alarm is generated.

Step 4 About one minute later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the maximum number of file objects allocated to the tenant is appropriate.

Step 5 On the FusionInsight Manager portal, choose the **Tenant Resources** page, select the tenant for which the alarm is generated, and click **Resources**. Check whether the maximum number of file objects configured for the tenant directory for which

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the alarm is generated is proper based on the actual service status of the tenant directory.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 On the **Resources** page, click **Modify** to modify or delete the maximum number of file objects configured for the tenant directory.

Step 7 About one minute later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 9 Select **HDFS** in the required cluster and **NodeAgent** under **Manager** from the **Service**.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 20 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.120 ALM-14026 Blocks on DataNode Exceed the Threshold

Alarm Description

The system checks the number of blocks on each DataNode every 30 seconds. This alarm is generated when the number of blocks on the DataNode exceeds the threshold.

If **Trigger Count** is **1** and the number of blocks on the DataNode is less than or equal to the threshold, this alarm is cleared. If **Trigger Count** is greater than **1** and the number of blocks on the DataNode is less than or equal to 90% of the threshold, this alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14026	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If this alarm is reported, there are too many blocks on the DataNode. In this case, data writing into the HDFS may fail due to insufficient disk space.

Possible Causes

- The alarm threshold is improperly configured.
- Data skew occurs among DataNodes.
- The disk space configured for the HDFS cluster is insufficient.

Handling Procedure

Change the threshold.

Step 1 On the FusionInsight Manager homepage, choose **Cluster** > **HDFS**. Then choose **Configurations** > **All Configurations**. On the displayed page, find the **GC_OPTS** parameter under **HDFS->DataNode**.

Step 2 Configure the threshold of the number of DataNodes.

Specifically, change the value of **Xmx** of the **GC_OPTS** parameter. **Xmx** specifies the memory, and each GB memory supports a maximum of 500,000 DataNode blocks. Set the memory as required. Confirm that **GC_PROFILE** is set to **custom** and save the configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Choose **Cluster** > **HDFS** > **Instance**. Select the DataNode instance whose status is **Expired**, click **More**, and select **Restart Instance** to make the **GC_OPTS** configuration take effect.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

Step 4 Check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether associated alarms are reported.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether the **ALM-14002 DataNode Disk Usage Exceeds the Threshold** alarm exists.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Handle the alarm by following the instructions in **ALM-14002 DataNode Disk Usage Exceeds the Threshold** and check whether the alarm is cleared.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 8**.

Expand the DataNode capacity.

Step 8 Expand the DataNode capacity.

Step 9 On FusionInsight Manager, wait for 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 20 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

Configuration rules of the DataNode JVM parameter.

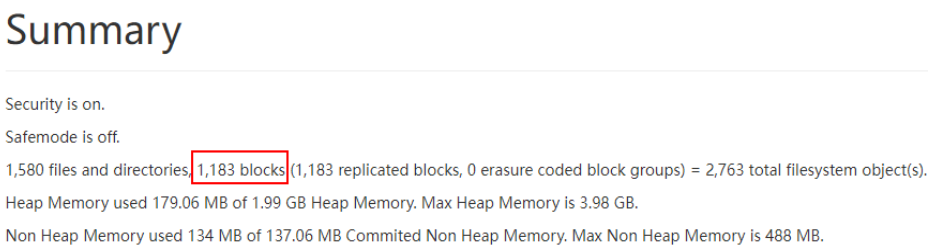
Default value of the DataNode JVM parameter **GC_OPTS**:

```
-Xms2G -Xmx4G -XX:NewSize=128M -XX:MaxNewSize=256M -XX:MetaspaceSize=128M -
XX:MaxMetaspaceSize=128M -XX:+UseConcMarkSweepGC -XX:+CMSParallelRemarkEnabled -
XX:CMSInitiatingOccupancyFraction=65 -XX:+PrintGCDetails -
Dsun.rmi.dgc.client.gcInterval=0x7FFFFFFFFFFFFE -Dsun.rmi.dgc.server.gcInterval=0x7FFFFFFFFFFFFE -XX:-
OmitStackTraceInFastThrow -XX:+PrintGCDateStamps -XX:+UseGCLogFileRotation -
XX:NumberOfGCLogFiles=10 -XX:GCLogFileSize=1M -Djdk.tls.ephemeralDHKeySize=2048
```

Average number of saved blocks = Number of HDFS blocks x 3/Number of DataNodes

To obtain the number of HDFS blocks, log in to FusionInsight Manager, choose **Cluster** > **Services** > **HDFS**, click **NameNode(*xxx*,Active)** to the right of **NameNode Web UI** to go to the native HDFS web UI, and then view the information in the **Summary** area.

Figure 9-7 Run the following command to view the number of HDFS blocks:



If the average number of blocks on a single DataNode has changed, modify **-Xms2G -Xmx4G -XX:NewSize=128M -XX:MaxNewSize=256M** in the default value. The following table lists the reference values.

Table 9-119 DataNode JVM configuration

Average Number of Blocks in a DataNode Instance	Reference Value
2,000,000	-Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M
5,000,000	-Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G

Xmx specifies memory which corresponds to the threshold of the number of DataNode blocks, and each GB memory supports a maximum of 500,000 DataNode blocks. Set the memory as required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.121 ALM-14027 DataNode Disk Fault

Alarm Description

The system checks the disk status on DataNodes every 60 seconds. This alarm is generated when a disk is faulty.

After all faulty disks on the DataNode are recovered, you need to manually clear the alarm and restart the DataNode.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14027	Major	No

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Failed Volumes	Specifies the list of faulty disks.

Impact on the System

If this alarm is reported, there are abnormal disk partitions on the DataNode. This may cause the loss of written files.

Possible Causes

- The hard disk is faulty.
- The disk permissions are configured improperly.

Handling Procedure

Check whether a disk alarm is generated.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and check whether **ALM-12014 Partition Lost** or **ALM-12033 Slow Disk Fault** exists.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Rectify the fault by referring to the handling procedure of **ALM-12014 Partition Lost** or **ALM-12033 Slow Disk Fault**. Then, check whether the alarm is cleared.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Modify disk permissions.

Step 4 Choose **O&M** > **Alarm** > **Alarms** and view **Location** and **Additional Information** of the alarm to obtain the location of the faulty disk.

Step 5 Log in to the node for which the alarm is generated as user **root**. Go to the directory where the faulty disk is located, and run the **ll** command to check whether the permission of the faulty disk is **711** and whether the user is **omm**.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Modify the permission of the faulty disk. For example, if the faulty disk is **data1**, run the following commands:

chown omm:wheel data1

chmod 711 data1

Step 7 In the alarm list on Manager, click **Clear** in the **Operation** column of the alarm to manually clear the alarm. Choose **Cluster** > **Services** > **HDFS** > **Instance**, select the DataNode, choose **More** > **Restart Instance**, wait for 5 minutes, and check whether a new alarm is reported.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

- If no, no further action is required.
- If yes, go to **Step 8**.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **HDFS** and **OMS** for the target cluster.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 20 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

After the fault is rectified, the system does not automatically clear this alarm and you need to manually clear the alarm.

Related Information

None

9.11.122 ALM-14028 Number of Blocks to Be Supplemented Exceeds the Threshold

Alarm Description

The system checks the number of blocks to be supplemented every 30 seconds and compares the number with the threshold. The number of blocks to be supplemented has a default threshold. This alarm is generated when the number of blocks to be supplemented exceeds the threshold.

To change the threshold, you can choose **O&M** > **Alarm** > **Thresholds** > **HDFS** > **File and Block** and click **Blocks Under Replicated (NameNode)**.

If **Trigger Count** is **1** and the number of blocks to be supplemented is less than or equal to the threshold, this alarm is cleared. If **Trigger Count** is greater than **1** and the number of blocks to be supplemented is less than or equal to 90% of the threshold, this alarm is cleared.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14028	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
NameServiceName	Specifies the NameService for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Data stored in HDFS is lost. HDFS may enter the security mode and cannot provide write services. Lost block data cannot be restored.

Possible Causes

- The DataNode instance is abnormal.
- Data is deleted.
- The number of replicas written into the file is greater than the number of DataNodes.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, check whether alarm **ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold** is generated.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Rectify the fault according to the handling procedure in **ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold**. Five minutes later, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Log in to the node where the MRS cluster client is installed, as user **root**. The user password is defined by the user before the installation. Contact the MRS cluster administrator to obtain the password.

1. If Kerberos authentication is not enabled for the MRS cluster, run the following command to switch to user **omm**:
   ```
   su - omm
   ```
2. Run the following command to go to the client installation directory and load environment variables:
   ```
   cd Client installation directory
   source bigdata_env
   ```
3. If Kerberos authentication is enabled for the cluster, complete security authentication:
   ```
   kinit Username of the service user who has the HDFS operation permissions
   ```
   Enter the required password.

Step 4 Run the following command to check integrity of the file system and obtain the status of the current cluster:

```
hdfs fsck / >> fsck.log
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Run the following command to count the number (*M*) of the blocks to be replicated:

```
cat fsck.log | grep "Under-replicated"
```

Step 6 Run the following command to count the number (*N*) of blocks to be replicated in the **/tmp/hadoop-yarn/staging/** directory:

```
cat fsck.log | grep "Under replicated" | grep "/tmp/hadoop-yarn/staging/" | wc -l
```

NOTE

/tmp/hadoop-yarn/staging/ is the default directory. If the directory is modified, obtain it from the configuration item **yarn.app.mapreduce.am.staging-dir** in the **mapred-site.xml** file.

Step 7 Check whether the percentage of *N* is greater than 50% (N/M > 50%).

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

Step 8 Run the following command to reconfigure the number of file replicas in the directory (set the number of file replicas to the number of DataNodes or the default number of file replicas):

```
hdfs dfs -setrep -w Number of file replicas/tmp/hadoop-yarn/staging/
```

Log in to FusionInsight Manager, choose **Cluster** > **Services** > **HDFS**. On the displayed page, click **Configurations** > **All Configurations**, and search for **dfs.replication**.

Step 9 Check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect the fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 12 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.123 ALM-14029 Number of Blocks in a Replica Exceeds the Threshold

Alarm Description

The system checks the number of blocks in a single replica every four hours and compares the number with the threshold. There is a threshold for the number of blocks in a single replica. This alarm is generated when the actual number of blocks in a single replica exceeds the threshold.

This alarm is cleared when the number of blocks to be supplemented is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14029	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
NameServiceName	Specifies the NameService for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Replica data is prone to be lost when a node is faulty. Too many files of a single replica affect the security of the HDFS file system.

Possible Causes

- The DataNode is faulty.
- The disk is faulty.
- Files are written to a single replica.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, check whether alarm **ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold** is generated.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Rectify the fault according to the handling procedure in **ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold**. In the next detection period, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Check whether files of a single replica have been written into the service.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 Log in to the HDFS client as user **root**. The user password is defined by the user before the installation. Contact the MRS cluster administrator to obtain the password. Run the following commands:

- Security mode:

  cd *Client installation directory*

  source bigdata_env

  kinit hdfs

- Normal mode:

  su - omm

  cd *Client installation directory*

  source bigdata_env

Step 5 Run the following command on the client node to increase the number of replicas for a single replica file:

hdfs dfs -setrep -w *file replica number file name or file path*

Step 6 In the next detection period, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.124 ALM-14030 HDFS Allows Write of Single-Replica Data

Alarm Description

This alarm is generated when **dfs.single.replication.enable** is set to **true**, indicating that HDFS is configured to allow write of single-replica data.

This alarm is cleared when this function is disabled on HDFS.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14030	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.

Impact on the System

If this configuration is enabled on the server and the number of HDFS replicas configured on the client is 1, single-replica data can be written to HDFS. Data of a single replica may be lost. Therefore, the system does not allow write of single-replica data by default. If a service requires single-replica data write to a directory, modify the HDFS configuration item **dfs.single.replication.exclude.pattern**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The HDFS configuration item **dfs.single.replication.enable** is set to **true**.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **HDFS**. On the page that is displayed, click the **Configurations** tab then the **All Configurations** sub-tab.

Step 2 Search for **dfs.single.replication.enable** in the search box, change the value of the configuration item to **false**, and click **Save**.

Step 3 Wait for about 10 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.125 ALM-14031 DataNode Process Is Abnormal

Alarm Description

The DataNode process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14031	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Then go to **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check the process state:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.hdfs.server.datanode.DataNode | grep -v grep | awk '{print $1}'

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.126 ALM-14032 JournalNode Process Is Abnormal

Alarm Description

The JournalNode process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14032	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Run the command in **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check the process state:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.hdfs.qjournal.server.JournalNode | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.127 ALM-14033 ZKFC Process Is Abnormal

Alarm Description

The ZKFC process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14033	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Go to **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check whether the process state is abnormal:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.hdfs.tools.DFSZKFailoverController | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.128 ALM-14034 Router Process Is Abnormal

Alarm Description

The Router process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14034	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Go to **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check whether the process state is abnormal:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.hdfs.server.federation.router.DFSRouter | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.129 ALM-14035 HttpFS Process Is Abnormal

Alarm Description

The HttpFS process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14035	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Go to **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check whether the process state is abnormal:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.fs.http.server.HttpFSServerWebServer | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.130 ALM-14036 NameNode Is In Safe Mode

Alarm Description

The system checks the NameNode process status every 30 seconds. This alarm is generated when the NameNode is in the safe mode.

This alarm is cleared when the process status recovers.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14036	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

After a NameNode enters the safe mode, data cannot be written to the NameNode.

Possible Causes

Block loss occurs when a user manually enters the safe mode or restarts the NameNode.

Handling Procedure

Check whether NameNode is in the security mode.

Step 1 Log in to FusionInsight Manager, click **O&M**, and choose **Alarm** > **Alarms** to view the alarm details. In the **Location** column, check the name of the host for which the alarm is generated.

Step 2 Choose **Cluster** > **Services** > **HDFS**, and click **NameNode(***host name recorded in **Step 1**,x***)** next to **NameNode Web UI**.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 3 In the **Basic Information** area on the **Dashboard** page of HDFS (or in the **NameService Summary** area on the **Dashboard** page of HDFS), check whether the value of **Safe Mode** is **ON**.

ON indicates that the safe mode is enabled.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 Log in to the HDFS client.

1. Log in to the node where the HDFS client is installed.
   - If Kerberos authentication is enabled for the cluster (the cluster is in security mode), log in as the **root** user.
   - If Kerberos authentication is disabled for the cluster (the cluster is in normal mode), log in as user **omm** and ensure that user **omm** has the execute permission on the client.

   (Note that you need to decide with cluster security/normal mode, not the HDFS safe/common mode.)
2. Run the following commands to go to the client installation directory and configure environment variables:

   cd *HDFS client installation directory*

   source bigdata_env
3. If Kerberos authentication is enabled for the cluster (the cluster is in security mode), run the following command to authenticate the user. If Kerberos authentication is disabled for the cluster (the cluster is in normal mode), skip this step.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

kinit hdfs

Enter the password as prompted. You can obtain the password from the MRS cluster administrator. Change the password upon the first authentication.

4. Run the following command to exit from the safe mode:

   hdfs dfsadmin -safemode leave

Step 5 Wait 5 minutes and check whether the alarm is cleared.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 In the **Basic Information** area on the **Dashboard** page of HDFS (or in the **NameService Summary** area on the **Dashboard** page of HDFS), check whether the value of **Missing Blocks** is **0**.

- If yes, no further action is required.
- If no, check whether "ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold" is reported and rectify the fault according to the alarm help.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.131 ALM-14037 DataNodes Outside the Cluster

Alarm Description

The NameNode checks whether there are DataNodes that are not managed in the cluster every 8 hours. This alarm is generated when there is a DataNode outside the cluster. This alarm is cleared when no DataNode is outside the cluster.

NOTE

This alarm applies only to MRS 3.5.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14037	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	NameService Name	Specifies the NameService for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition, that is, the IP address and port of a DataNode outside the cluster is detected.

Impact on the System

Data may be lost.

Possible Causes

After a host is forcibly deleted, the host is powered on again, and the process is restarted.

Handling Procedure

Step 1 Log in to FusionInsight Manager, click **O&M**, and choose **Alarm** > **Alarms** to view the alarm details. In the additional information area, check the IP address of the host for which the alarm is generated.

Step 2 Stop the DataNode process on the host for which the alarm is reported.

NOTICE

If there are multiple IP addresses of the host, you can **stop only one DataNode process at a time** and stop the next DataNode process only after **Number of Blocks to Be Replicated** changes to **0**.

1. Log in to the host for which the alarm is generated as the **root** user and change the permission on the **hadoop** directory in the installation directory **${BIGDATA_HOME}/FusionInsight_HD_*/install**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

chmod 000 ${BIGDATA_HOME}/FusionInsight_HD_*xxx*/install/FusionInsight-Hadoop-3.3.1

2. Run the following commands to obtain the PID of the DataNode process and stop it on the host:

   ps -ef | grep Dproc_datanode

   kill -15 *PID*

3. Choose **Cluster** > **Services** > **HDFS**. Check the **Basic Information** area in the **Dashboard** tab (or the **NameService Summary** area in the **Dashboard** tab of HDFS), and wait until the value of **Blocks to be Replicated** changes to **0**.

Step 3 Wait for 8 hours and check whether the alarm is cleared and whether the number of blocks to be replicated is 0.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HDFS** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

9.11.132 ALM-14038 Router Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the size of the used HDFS Router heap memory and the maximum size of the heap memory that can be allocated every 30 seconds, calculates the ratio of the used heap memory to the maximum size of the heap memory that can be allocated to obtain the heap memory usage, and compares the actual heap memory usage of the HDFS Router with the threshold. The HDFS Router Heap Memory usage has a default threshold. This alarm is generated when the HDFS Router Heap Memory usage exceeds the threshold.

You can change the threshold in **O&M** > **Alarm > Thresholds** > *Name of the desired cluster* **> HDFS**.

The alarm is cleared when the heap memory usage is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14038	Critical (default threshold: 95%) Major (default threshold: 90%)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm is generated.
	ServiceName	Specifies the service for which the alarm is generated.
	RoleName	Specifies the role for which the alarm is generated.
	HostName	Specifies the host for which the alarm is generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

The HDFS Router Heap Memory usage is too high, which affects the data read/write performance of the HDFS.

Possible Causes

The HDFS Router Heap Memory is insufficient.

Handling Procedure

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **PolicySync**. Click **Instance Configuration** and then **All Configurations**, and choose **PolicySync** > **System**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **HDFS** > **Configurations** > **All Configurations**. In **Search**, enter **GC_OPTS** to check the GC_OPTS memory parameter of **HDFS->Router**.

Step 3 Increase the values of **-Xms** and **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

If this alarm is generated, the heap memory configured for Router cannot meet the requirements of the current process. You are advised to change the values of **-Xms** and **-Xmx** in the **GC_OPTS** parameter to twice the size of the used heap memory or change the values based on the site requirements.

Step 4 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **HDFS** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.133 ALM-14039 Slow DataNodes Exist in the Cluster

Alarm Description

The system checks the number of slow operations per second on HDFS DataNode instances every 60 seconds and compares the number with the threshold. This alarm is generated when the number of slow operations per second of an HDFS DataNode instance has exceeded the threshold for three minutes.

This alarm is cleared when the number of slow operations per second of the HDFS DataNode instance is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
14039	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

Slow DataNodes on HDFS affect the data read and write performance of HDFS.

Possible Causes

- The disk I/O rate of the HDFS DataNode instance is low, and the HDFS DataNode processing capability reaches the bottleneck.
- The network transmission rate between HDFS DataNode instances is low.

Handling Procedure

Check whether the disk I/O rate of the DataNode instance is low.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the host name of the DataNode instance for which this alarm is generated.

Step 2 Choose **Cluster** > **Services** > **HDFS**, click the **Instances** tab, and click the DataNode role based on the host name obtained in **Step 1**.

Step 3 Click the **Chart** tab and select **Performance** from the **Chart Category** area. Check whether any data in **Slow Flush or Sync Occurrences Per Second**, **Slow SyncWriterOsCache Occurrences Per Second**, and **Slow WriteDataToDisk Occurrences Per Second** charts is high.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and check whether **ALM-12033 Slow Disk Fault** exists.

- If yes, record the disk information in the alarm details and go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Obtain information about the disk where slow operations occur.

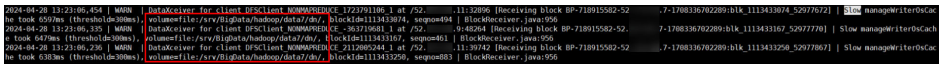
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to the DataNode using the IP address obtained in **Step 1** as user **omm** and run the following commands to view the run log:

   cd /var/log/Bigdata/hdfs/dn/

   vim hadoop-omm-datanode-*Hostname*.log

2. Search for keyword **slow** in the log to identify the disk where slow operations occur.



Step 6 Rectify the fault based on the obtained disk information by following the handling procedure of **ALM-12033 Slow Disk Fault**.

Step 7 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the network transmission rate between HDFS DataNode instances is low.

Step 8 On FusionInsight Manager, choose **Cluster** > **Services** > **HDFS**, click the **Chart** tab, select **Performance** in the **Chart Category** area, and check whether any data in the **Slow Write Packet To DownStream Count Per Second** and **Slow Ack To Upstream Count Per Second** charts is high.

- If yes, go to **Step 9**.
- If no, go to **Step 13**.

Step 9 Log in to the DataNode using the IP address obtained in **Step 1** as user **omm** and run the following commands to view the run log:

cd /var/log/Bigdata/hdfs/dn/

vim hadoop-omm-datanode-*Hostname*.log

Step 10 Search for keyword **slow** in the log to identify the upstream and downstream nodes where slow operations occur.



Step 11 Check whether the network communication between the current node and the nodes obtained in **Step 10** is normal.

- If yes, go to **Step 13**.
- If no, contact the network administrator to repair the network.

Step 12 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **HDFS** for the target cluster.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.134 ALM-16000 Percentage of Sessions Connected to the HiveServer to Maximum Number Allowed Exceeds the Threshold

Description

The system detects the percentage of sessions connected to the HiveServer to the maximum number of allowed sessions every 30 seconds. This indicator can be viewed on the **Cluster** > *Name of the desired cluster* > **Services** > **Hive > Instance** > *HiveServer instance***.** This alarm is generated when the the percentage of sessions connected to the HiveServer to the maximum number allowed exceeds the specified threshold (90% by default).

To change the threshold, choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* > **Hive > Percentage of Sessions Connected to the HiveServer to Maximum Number of Sessions Allowed by the HiveServer**.

When the **Trigger Count** is 1, this alarm is cleared when the percentage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the percentage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
16000	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If a connection number alarm is generated, the number of sessions connected to HiveServer is too large. As a result, new connections cannot be established, new tasks fail, or even services restart unexpectedly.

Possible Causes

Too many clients are connected to HiveServer.

Procedure

Increase the maximum number of connections to Hive.

Step 1 On the FusionInsight Manager portal, Choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Configurations > All Configurations**.

Step 2 Search for **hive.server.session.control.maxconnections** and increase the value of this parameter. If the value of this parameter is **A**, the threshold is **B**, and the number of sessions connected to the HiveServer is **C**, adjust the value of this parameter according to **A x B > C**. To view the number of sessions connected to the HiveServer, check the value of **Statistics for Sessions of the HiveServer** on the Hive monitoring page.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 5 Select **Hive** in the required cluster from the **Service**.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.135 ALM-16001 Hive Warehouse Space Usage Exceeds the Threshold

Description

The system checks the Hive warehouse space usage every 30 seconds. The metric **Percentage of HDFS Space Used by Hive to the Available Space** can be viewed on the Hive service monitoring page. This alarm is generated when the Hive warehouse space usage exceeds the specified threshold (85% by default).

To change the threshold, choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* **> Hive > Percentage of HDFS Space Used by Hive to the Available Space**.

When the **Trigger Count** is 1, this alarm is cleared when the Hive warehouse space usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the Hive warehouse space usage is less than or equal to 90% of the threshold.

NOTE

The MRS cluster administrator can reduce the repository space usage by increasing the repository capacity or releasing some used space.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
16001	Minor	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The system cannot write data properly. Some data may be lost.

Possible Causes

- The upper limit of the HDFS capacity available for Hive is too small.
- The HDFS space is insufficient.
- Some data nodes break down.

Procedure

Expand the system configuration.

Step 1 Analyze the cluster HDFS space usage and increase the HDFS capacity for Hive.

Log in to FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Configuration**, select **All Configurations**, search for **hive.metastore.warehouse.size.percent**, and increase the value. Assume that the value of the configuration item is A, the total HDFS storage space is B, the threshold is C, and the HDFS space used by Hive is D. Adjust the value by complying with A x B x C > D. You can view the total HDFS storage space on the HDFS NameNode page, and the HDFS space used by Hive on the Hive monitoring page.

Step 2 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Expand the system.

Step 3 Expand the system.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the data node is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 On the FusionInsight Manager portal, click **O&M > Alarm > Alarms**.

Step 6 Check whether "ALM-12006 Node Fault", "ALM-12007 Process Fault", or "ALM-14002 DataNode Disk Usage Exceeds the Threshold" exist.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Handle the alarms by following the handling procedures in **ALM-12006 Node Fault**, **ALM-12007 Process Fault**, and **ALM-14002 DataNode Disk Usage Exceeds the Threshold**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 10 Select **Hive** in the required cluster from the **Service**.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.136 ALM-16002 Hive SQL Execution Success Rate Is Lower Than the Threshold

Description

The system checks the percentage of the HQL statements that are executed successfully in every 30 seconds. The formula is: Percentage of HQL statements that are executed successfully = Number of HQL statements that are executed successfully by Hive in a specified period/Total number of HQL statements that are executed by Hive. This indicator can be viewed on the **Cluster >** *Name of the desired cluster* **> Services** > **Hive > Instance** > *HiveServer instance*. The default threshold of the percentage of HQL statements that are executed successfully is **90%**. An alarm is reported when the percentage is lower than the **90%**. Users can view the name of the host where an alarm is generated in the location information about the alarm. The IP address of the host is the IP address of the HiveServer node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Users can modify the threshold by choosing **O&M > Alarm > Thresholds >** *Name of the desired cluster* > **Hive** > **Percentage of HQL Statements That Are Executed Successfully by Hive**.

This alarm is cleared when the execution success rate is higher than 110% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
16002	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The service execution capability of the system is too low and cannot properly respond to customer requests. The Hive service is not affected. You need to check HiveServer logs to locate the SQL failure cause.

Possible Causes

- A syntax error occurs in HQL statements.
- The HBase service is abnormal when a Hive on HBase task is performed.
- The Spark service is abnormal when a Hive on Spark task is performed.
- The dependent basic services, such as HDFS, Yarn, and ZooKeeper, are abnormal.

Procedure

Check whether the HQL statements comply with syntax.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On the FusionInsight Manager page, choose **O&M** > **Alarm** to view the alarm details and obtain the node where the alarm is generated.

Step 2 Use the Hive client to log in to the HiveServer node where an alarm is reported. Query the HQL syntax provided by Apache, and check whether the HQL commands are correct.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

NOTE

To view the user who runs an incorrect statement, you can download the HiveServerAudit log file of the HiveServer node where this alarm is generated. **Start Data** and **End Data** are 10 minutes before and after the alarm generation time respectively. Open the log file and search for the **Result=FAIL** keyword to filter the log information about the incorrect statement, and then view the user who runs the incorrect statement according to **UserName** in the log information.

Step 3 Enter the correct HQL statements, and check whether the command can be properly executed.

- If yes, go to **Step 12**.
- If no, go to **Step 4**.

Check whether the HBase service is abnormal.

Step 4 Check whether a Hive on HBase task is performed with the user who runs the HQL command.

- If yes, go to **Step 5**.
- If no, go to **Step 8**.

Step 5 On the FusionInsight Manager page, click **Cluster** > *Name of the desired cluster* > **Services**, check whether the HBase service is normal in the service list.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Choose **O&M** > **Alarm**, check the related alarms displayed on the alarm page and clear them according to related alarm help.

Step 7 Enter the correct HQL statements, and check whether the command can be properly executed.

- If yes, go to **Step 12**.
- If no, go to **Step 8**.

Check whether the HDFS, Yarn, and ZooKeeper are normal.

Step 8 On the FusionInsight Manager portal, click **Cluster** > *Name of the desired cluster* > **Services**.

Step 9 In the service list, check whether the services, such as HDFS, Yarn, and ZooKeeper are normal.

- If yes, go to **Step 12**.
- If no, go to **Step 10**.

Step 10 Check the related alarms displayed on the alarm page and clear them according to related alarm help.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 Enter the correct HQL statements, and check whether the command can be properly executed.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 After 1 minute, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On the FusionInsight Manager home page, choose **O&M** > **Log** > **Download**.

Step 14 Select the following nodes in the required cluster from the **Service**:

- MapReduce
- Hive

Step 15 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.137 ALM-16003 Background Thread Usage Exceeds the Threshold

Description

The system checks the usage of the background thread every 30 seconds. The default threshold is **90%**. This alarm is generated when the usage of the background thread pool of Hive exceeds the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16003	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

There are too many background threads, so the newly submitted task cannot run in time.

Possible Causes

The usage of the background thread pool of Hive is excessively high when:

- There are many tasks executed in the background thread pool of HiveServer.
- The capacity of the background thread pool of HiveServer is too small.

Procedure

Check the number of tasks executed in the background thread pool of HiveServer.

Step 1 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive**. On the displayed page, click **HiveServer Instance** and check values of **Background Thread Count** and **Background Thread Usage**.

Step 2 Check whether the number of background threads in the latest half an hour is excessively high. (By default, the queue number is 100, and the thread number is considered as high if it is 90 or larger.)

- If it is, go to **Step 3**.
- If it is not, go to **Step 5**.

Step 3 Adjust the number of tasks submitted to the background thread pool. (For example, cancel some time-consuming tasks with low performance.)

Step 4 Check whether the values of Background Thread Count and Background Thread Usage decrease.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is, go to **Step 7**.
- If it is not, go to **Step 5**.

Check the capacity of the HiveServer background thread pool.

Step 5 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive**. On the displayed page, click **HiveServer Instance** and check values of **Background Thread Count** and **Background Thread Usage**.

Step 6 Increase the value of **hive.server2.async.exec.threads** in the **${BIGDATA_HOME}/FusionInsight_HD_8.1.0.1/1_23_HiveServer/etc/hive-site.xml** file. For example, increase the value by 20%.

Step 7 Save the modification.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 10 Select **Hive** in the required cluster from the **Service**.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.138 ALM-16004 Hive Service Unavailable

Description

The system checks the Hive service status every 60 seconds. This alarm is generated when the Hive service is unavailable.

This alarm is cleared when the Hive service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
16004	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The system cannot provide data loading, query, and extraction services.

Possible Causes

- Hive service unavailability may be related to the faults of the Hive process as well as basic services, such as ZooKeeper, Hadoop distributed file system (HDFS), YARN, and DBService.
  - The ZooKeeper service is abnormal.
  - The HDFS service is abnormal.
  - The YARN service is abnormal.
  - The DBService service is abnormal.
  - The Hive service process is abnormal. If the alarm is caused by Hive process fault, the alarm report has a delay of about 5 minutes.
- The network communication between the Hive and basic services is interrupted.
- The permission on the HDFS temporary directory of Hive is abnormal.
- The local disk space of the Hive node is insufficient.

Procedure

Check the HiveServer/MetaStore process status.

Step 1 On the FusionInsight Manager portal, click **Cluster >** *Name of the desired cluster* **> Services** > **Hive** > **Instance**. In the Hive instance list, check whether the HiveServer or MetaStore instances are in the Unknown state.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 In the Hive instance list, choose **More** > **Restart Instance** to restart the HiveServer/MetaStore process.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

During HiveServer or MetaStore instance restart, the instance cannot provide services for external systems. SQL tasks that are being executed on the instance may fail.

Step 3 In the alarm list, check whether **Hive Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the ZooKeeper service status.

Step 4 On the FusionInsight Manager, check whether the alarm list contains **Process Fault**.

- If yes, go to **Step 5**.
- If no, go to **Step 8**.

Step 5 In the **Process Fault**, check whether **ServiceName** is **ZooKeeper**.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Rectify the fault by following the steps provided in **ALM-12007 Process Fault**.

Step 7 In the alarm list, check whether **Hive Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check the HDFS service status.

Step 8 On the FusionInsight Manager, check whether the alarm list contains **HDFS Service Unavailable**.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Rectify the fault by following the steps provided in **ALM-14000 HDFS Service Unavailable**.

Step 10 In the alarm list, check whether **Hive Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check the Yarn service status.

Step 11 In FusionInsight Manager alarm list, check whether **Yarn Service Unavailable** is generated.

- If yes, go to **Step 12**.
- If no, go to **Step 14**.

Step 12 Rectify the fault by performing the operations provided for **ALM-18000 Yarn Service Unavailable**.

Step 13 In the alarm list, check whether **Hive Service Unavailable** is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 14**.

Check the DBService service status.

Step 14 In FusionInsight Manager alarm list, check whether **DBService Service Unavailable** is generated.

- If yes, go to **Step 15**.
- If no, go to **Step 17**.

Step 15 Rectify the fault. For details, see **ALM-27001 DBService Service Unavailable**.

Step 16 In the alarm list, check whether **Hive Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 17**.

Check the network connection between the Hive and ZooKeeper, HDFS, Yarn, and DBService.

Step 17 On the FusionInsight Manager, choose **Cluster >** *Name of the desired cluster* **> Services** > **Hive**.

Step 18 Click **Instance**.

The HiveServer instance list is displayed.

Step 19 Click **Host Name** in the row of **HiveServer**.

The active HiveServer host status page is displayed.

Step 20 Record the IP address under **Basic Information**.

Step 21 Use the IP address obtained in **Step 20** to log in to the host where the active HiveServer runs as user **omm**.

Step 22 Run the **ping** command to check whether communication between the host that runs the active HiveServer and the hosts that run the ZooKeeper, HDFS, Yarn, and DBService services are normal. (Obtain the IP addresses of the hosts that run the ZooKeeper, HDFS, Yarn, and DBService services in the same way as that for obtaining the IP address of the active HiveServer.)

- If yes, go to **Step 25**.
- If no, go to **Step 23**.

Step 23 Contact the administrator to restore the network.

Step 24 In the alarm list, check whether **Hive Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 25**.

Check the permission on the HDFS temporary directory.

Step 25 Log in to the node where the HDFS client is located and run the following command to go to the HDFS client installation directory:

cd *Client installation directory*

source bigdata_env

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

kinit *user with the supergroup permission* (Skip this step for common clusters.)

Step 26 Run the following command to check whether the permission on the data warehouse directory is 770:

hdfs dfs -ls /tmp | grep hive-scratch

- If yes, go to **Step 29**.
- If no, go to **Step 27**.

Step 27 Run the following command to restore the default data warehouse permission:

hdfs dfs -chmod 770 /tmp/hive-scratch

Step 28 Wait for several minutes and check whether the Hive Service Unavailable alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 29**.

Check whether the local disk space is normal.

Step 29 Run the **df -h** command to check the root directory and check whether the disk usage of the **/srv**, **/var**, and **/opt** directories exceeds 95%.

- If yes, go to **Step 30**.
- If no, go to **Step 31**.

Step 30 Clear unnecessary information in the corresponding directory to ensure that the available disk space is greater than 80%. Wait for several minutes and check whether the Hive Service Unavailable alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 31**.

Collect fault information.

Step 31 On the FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 32 Select the following nodes in the required cluster from the **Service**:

- ZooKeeper
- HDFS
- Yarn
- DBService
- Hive

Step 33 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 34 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.139 ALM-16005 The Heap Memory Usage of the Hive Process Exceeds the Threshold

Alarm Description

The system checks the Hive service status every 30 seconds. The alarm is generated when the heap memory usage of a Hive service exceeds the threshold (95% of the maximum memory).

You can choose **O&M** > **Alarm** > **Thresholds** > **Hive** and change the threshold.

The alarm is cleared when the Hive heap memory usage is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
16005	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

When the heap memory usage of Hive is overhigh, the performance of Hive task operation is affected. In addition, a memory overflow may occur so that the Hive service is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The heap memory of the Hive instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Handling Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose **Alarm ID** is **16005**. Then check the role name in **Location** and confirm the IP address of the instance.

- If the role for which the alarm is generated is HiveServer, go to **Step 2**.
- If the role for which the alarm is generated is MetaStore, go to **Step 3**.

Step 2 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Hive** > **Instances** and click the HiveServer for which the alarm was generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area, choose **Customize** > **CPU and Memory**, select **HiveServer Memory Usage Statistics**, and click **OK**. Check whether the used heap memory of the HiveServer service reaches the threshold (default value: 95%) of the maximum heap memory specified for HiveServer.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Hive** > **Instances** and click the MetaStore for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory**, and select **MetaStore Memory Usage Statistics** and click **OK**. Check whether the used heap memory of the MetaStore service reaches the threshold(default value: 95%) of the maximum heap memory specified for MetaStore.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Hive** > **Configurations** > **All Configurations**. Choose **HiveServer/MetaStore** > **JVM**. Adjust the value of **-Xmx** in **HIVE_GC_OPTS/METASTORE_GC_OPTS** according to the following rules. Click **Save**.

1. Suggestions for GC parameter settings for the HiveServer:

   When the heap memory used by the HiveServer process reaches the heap memory threshold set for the HiveServer process, change the value of **-Xmx** to twice the default value. For example, if **-Xmx** is set to 2 GB by default, change the value of **-Xmx** to 4 GB.

   On the FusionInsight Manager homepage, you can choose **O&M** > **Alarm** > **Thresholds** > **Hive** > **CPU and Memory** > **HiveServer Heap Memory Usage Statistics (HiveServer)** to view the threshold.

   Change the value of -Xms so that the ratio of -Xms to -Xmx is 1:2, which prevents performance from being affected when JVM dynamically adjusts the heap memory size.

2. Suggestions for GC parameter settings for the MetaServer:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

When the heap memory used by the MetaStore process reaches the heap memory threshold set for the MetaStore process, change the value of **-Xmx** to twice the default value. For example, if **-Xmx** is set to 2 GB by default, change the value of **-Xmx** to 4 GB.

On the FusionInsight Manager homepage, you can choose **O&M** > **Alarm** > **Thresholds** > **Hive** > **CPU and Memory** > **MetaStore Heap Memory Usage Statistics (MetaStore)** to view the threshold.

Change the value of -Xms so that the ratio of -Xms to -Xmx is 1:2, which prevents performance from being affected when JVM dynamically adjusts the heap memory size.

Step 5 Click **More > Restart Service** to restart the Hive service.

WARNING

During Hive service restart, instances cannot provide services for external systems, and the SQL tasks that are being executed on the instances may fail.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **Hive** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.140 ALM-16006 The Direct Memory Usage of the Hive Process Exceeds the Threshold

Description

The system checks the Hive service status every 30 seconds. The alarm is generated when the direct memory usage of a Hive service exceeds the threshold (95% of the maximum memory).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Users can choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* **> Hive** to change the threshold.

The alarm is cleared when the direct memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
16006	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the direct memory usage of Hive is overhigh, the performance of Hive task operation is affected. In addition, a memory overflow may occur so that the Hive service is unavailable.

Possible Causes

The direct memory of the Hive instance on the node is overused or the direct memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check direct memory usage.

Step 1 On the FusionInsight Manager portal, click **O&M > Alarm > Alarms** and select the alarm whose **Alarm ID** is **16006**. Then check the role name in **Location** and confirm the IP address of the instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the role for which the alarm is generated is HiveServer, go to **Step 2**.
- If the role for which the alarm is generated is MetaStore, go to **Step 3**.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services > Hive > Instance** and click the HiveServer for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory**, and select **HiveServer Memory Usage Statistics** and click **OK**, check whether the used direct memory of the HiveServer service reaches the threshold(default value: 95%) of the maximum direct memory specified for HiveServer.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Instance** and click the MetaStore for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory**, and select **MetaStore Memory Usage Statistics** and click **OK**, check whether the used direct memory of the MetaStore service reaches the threshold(default value: 95%) of the maximum direct memory specified for MetaStore.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Configurations > All Configurations**. Choose **HiveServer/MetaStore** > **JVM**. Adjust the value of **-XX:MaxDirectMemorySize** in **HIVE_GC_OPTS/METASTORE_GC_OPTS** as the following rules. Click **Save**.

NOTE

1. Suggestions for GC parameter settings for the HiveServer:
   - It is recommended that you set the value of **-XX:MaxDirectMemorySize** to 1/8 of the value of **-Xmx**. For example, if **-Xmx** is set to 8 GB, **-XX:MaxDirectMemorySize** is set to 1024 MB. If **-Xmx** is set to 4 GB, **-XX:MaxDirectMemorySize** is set to 512 MB. It is recommended that the value of **-XX:MaxDirectMemorySize** be greater than or equal to 512 MB.
2. Suggestions for GC parameter settings for the MetaServer:
   - It is recommended that you set the value of **-XX:MaxDirectMemorySize** to 1/8 of the value of **-Xmx**. For example, if **-Xmx** is set to 8 GB, **-XX:MaxDirectMemorySize** is set to 1024 MB. If **-Xmx** is set to 4 GB, **-XX:MaxDirectMemorySize** is set to 512 MB. It is recommended that the value of **-XX:MaxDirectMemorySize** be greater than or equal to 512 MB.

Step 5 Click **More > Restart Service** to restart the service.

WARNING

During Hive service restart, instances cannot provide services for external systems, and the SQL tasks that are being executed on the instances may fail.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Hive** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.141 ALM-16007 Hive GC Time Exceeds the Threshold

Alarm Description

The system checks the garbage collection (GC) time of the Hive service every 60 seconds. This alarm is generated when the detected GC time exceeds the threshold or exceeds 12 seconds for three consecutive checks. You can choose **O&M** > **Alarm** > **Thresholds** > **Hive** and change the threshold. This alarm is cleared when the Hive GC time is shorter than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
16007	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the GC time exceeds the threshold, Hive data read and write are affected, task execution may slow down, or services may restart unexpectedly.

Possible Causes

The memory of Hive instances is overused, the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Handling Procedure

Check the GC time.

Step 1 On the FusionInsight Manager homepage, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose **Alarm ID** is **16007**. Then check the role name in **Location** and confirm the IP address of the instance.

- If the role for which the alarm is generated is HiveServer, go to **Step 2**.
- If the role for which the alarm is generated is MetaStore, go to **Step 3**.

Step 2 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Hive** > **Instances** and click the HiveServer for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **GC**, and select **Garbage Collection (GC) Time of HiveServer** and click **OK** to check whether the GC time is longer than 12 seconds.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Hive** > **Instances** and click the MetaStore for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **GC**, and select **Garbage Collection (GC) Time of MetaStore** and click **OK** to check whether the GC time is longer than 12 seconds.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Check the current JVM configuration.

Step 4 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Hive** > **Configurations** > **All Configurations**. Choose **HiveServer/MetaStore** > **JVM**. Adjust the value of **-Xmx** in **HIVE_GC_OPTS/METASTORE_GC_OPTS** according to the following rules. Click **Save**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- Suggestions for GC parameter settings for the HiveServer:
  - When the Hive GC time exceeds the threshold, change the value of **-Xmx** to twice the default value. For example, if **-Xmx** is set to 2 GB by default, change the value of **-Xmx** to 4 GB.
  - You are advised to change the value of **-Xms** so that the ratio of **-Xms** to **-Xmx** is 1:2, which prevents performance from being affected when the JVM dynamically adjusts the heap memory size.
- Suggestions for GC parameter settings for the MetaServer:
  - When the Meta GC time exceeds the threshold, change the value of **-Xmx** to twice the default value. For example, if **-Xmx** is set to 2 GB by default, change the value of **-Xmx** to 4 GB.
  - You are advised to change the value of **-Xms** so that the ratio of **-Xms** to **-Xmx** is 1:2, which prevents performance from being affected when the JVM dynamically adjusts the heap memory size.

Step 5 Click **More** > **Restart Service** to restart the service.

WARNING

During Hive service restart, instances cannot provide services for external systems, and the SQL tasks that are being executed on the instances may fail.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal of active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 8 In the **Service**, select **Hive** in the required cluster.

Step 9 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.142 ALM-16008 Non-Heap Memory Usage of the Hive Process Exceeds the Threshold

Description

The system checks the Hive service status every 30 seconds. The alarm is generated when the non-heap memory usage of a Hive service exceeds the threshold (95% of the maximum memory).

Users can choose **O&M > Alarm > Thresholds >** *Name of the desired cluster* > **Hive** to change the threshold.

The alarm is cleared when the non-heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
16008	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

When the non-heap memory usage of Hive is overhigh, the performance of Hive task operation is affected. In addition, a memory overflow may occur so that the Hive service is unavailable.

Possible Causes

The non-heap memory of the Hive instance on the node is overused or the non-heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check non-heap memory usage.

Step 1 On the FusionInsight Manager portal, click **O&M > Alarm > Alarms** and select the alarm whose **Alarm ID** is **16008**. Then check the role name in **Location** and confirm the IP address of the instance.

- If the role for which the alarm is generated is HiveServer, go to **Step 2**.
- If the role for which the alarm is generated is MetaStore, go to **Step 3**.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Instance** and click the HiveServer for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory**, and select **HiveServer Memory Usage Statistics** and click **OK**, check whether the used non-heap memory of the HiveServer service reaches the threshold(default value: 95%) of the maximum non-heap memory specified for HiveServer.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Instance** and click the MetaStore for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory**, and select **MetaStore Memory Usage Statistics** and click **OK**, check whether the used non-heap memory of the MetaStore service reaches the threshold(default value: 95%) of the maximum non-heap memory specified for MetaStore.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* > **Services** > **Hive** > **Configurations > All Configurations**. Choose **HiveServer/MetaStore** > **JVM**. Adjust the value of **-XX:MaxMetaspaceSize** in **HIVE_GC_OPTS/METASTORE_GC_OPTS** as the following rules. Click **Save**.

NOTE

1. Suggestions for GC parameter settings for the HiveServer:
   - It is recommended that you set the value of **-XX:MaxMetaspaceSize** to 1/8 of the value of **-Xmx**. For example, if **-Xmx** is set to 2 GB, **-XX:MaxMetaspaceSize** is set to 256 MB. If **-Xmx** is set to 4 GB, **-XX:MaxMetaspaceSize** is set to 512 MB.
2. Suggestions for GC parameter settings for the MetaServer:
   - It is recommended that you set the value of **-XX:MaxMetaspaceSize** to 1/8 of the value of **-Xmx**. For example, if **-Xmx** is set to 2 GB, **-XX:MaxMetaspaceSize** is set to 256 MB. If **-Xmx** is set to 4 GB, **-XX:MaxMetaspaceSize** is set to 512 MB

Step 5 Click **More > Restart Service** to restart the service.

WARNING

During Hive service restart, instances cannot provide services for external systems, and the SQL tasks that are being executed on the instances may fail.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Hive** in the required cluster from the **Service**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.143 ALM-16009 Map Number Exceeds the Threshold

Description

The system checks the number of HQL maps in every 30 seconds. This alarm is generated if the number exceeds the threshold. By default, **Trigger Count** is set to **3**, and the threshold is 5000.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16009	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the number of HQL maps executed by Hive is too large, a large number of Yarn queue resources are occupied, which may take a long time and affect other tasks running using this queue.

Possible Causes

The HQL statements are not the optimal.

Procedure

Check the number of HQL maps.

Step 1 On FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Hive** > **Resource**. Check the HQL statements with the excessively large number (5000 or more) of maps in **HQL Map Count**.

Step 2 Locate the corresponding HQL statements, optimize them and execute them again.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If it is not, go to **Step 4**.

Collect fault information.

Step 4 On the FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 5 Select **Hive** in the required cluster from the **Service**.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.144 ALM-16045 Hive Data Warehouse Is Deleted

Description

The system checks the Hive data warehouse in every 60 seconds. This alarm is generated when the Hive data warehouse is deleted.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16045	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The default Hive data warehouse is deleted. As a result, creating databases or tables in the default data warehouse fails, and services are affected.

Possible Causes

Hive periodically checks the status of the default data warehouse and finds that the default data warehouse is deleted.

Procedure

Check the default Hive data warehouse.

Step 1 Log in to the node where the client is located as user **root**.

Step 2 Run the following command to check whether the **warehouse** directory exists in **hdfs://hacluster/user/***<username>***/.Trash/Current/**.

hdfs dfs -ls hdfs://hacluster/user/*<username>*/.Trash/Current/

For example, if **user/hive/warehouse** exists:

```
host01:/opt/client # hdfs dfs -ls hdfs://hacluster/user/test/.Trash/Current/
Found 1 items
drwx------   - test hadoop          0 2019-06-17 19:53 hdfs://hacluster/user/test/.Trash/Current/user
```

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 By default, there is an automatic recovery mechanism for the data warehouse. You can wait for 5 ~10s to check whether the default data warehouse is restored. If the data warehouse is not recovered, manually run the following command to restore the data warehouse.

hdfs dfs -mv hdfs://hacluster/user/*<username>*/.Trash/Current/user/hive/warehouse /user/hive/warehouse

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 Collect related information in the **.Trash/Current/** directory on the client background.

Step 6 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.145 ALM-16046 Hive Data Warehouse Permission Is Modified

Description

The system checks the Hive data warehouse permission in every 60 seconds. This alarm is generated if the permission is modified.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16046	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the permission on the Hive default data warehouse is modified, the permission for users or user groups to create databases or tables in the default data warehouse is changed.

Possible Causes

Hive periodically checks the status of the default data warehouse and finds that default data warehouse permission is changed.

Procedure

Check the Hive default data warehouse permission.

Step 1 Log in to the node where the client is located as user **root**.

Step 2 Run the following command to go to the HDFS client installation directory:

cd *Client installation directory*

source bigdata_env

kinit *User who has the supergroup permission* (Skip this step for a common cluster.)

Step 3 Run the following command to restore the default data warehouse permission:

- Security mode: **hdfs dfs -chmod 770 hdfs://hacluster/user/hive/warehouse**
- Non-security mode: **hdfs dfs -chmod 777 hdfs://hacluster/user/hive/warehouse**

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If it is not, go to **Step 5**.

Collect fault information.

Step 5 Collect related information in the **hdfs://hacluster/user/hive/warehouse** directory on the client background.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.146 ALM-16047 HiveServer Has Been Deregistered from ZooKeeper

Alarm Description

The system checks the Hive service every 60 seconds. This alarm is generated when Hive registration information on ZooKeeper is lost or Hive cannot connect to ZooKeeper.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
16047	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

When a Hive client sets up a new connection, it cannot select the HiveServer node that has been deregistered from ZooKeeper. If all HiveServer nodes have been deregistered from ZooKeeper, the HiveServer service will be unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The ZooKeeper instance is abnormal.
- Some Hive configurations are incorrect.

Handling Procedure

Check the ZooKeeper service status.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and check whether **ALM-12007 Process Fault** exists in the alarm list.

- If yes, go to **Step 2**.
- If no, go to **Step 5**.

Step 2 In **Location** of **ALM-12007 Process Fault**, check whether the service name is **ZooKeeper**.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by following the steps provided in **ALM-12007 Process Fault**.

Step 4 In the alarm list, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the Hive configurations are correctly modified.

Step 5 On FusionInsight Manager, choose **Audit**. On the **Audit** page, click **Advanced Search**, click ··· on the right of **Operation Type**, select **Save configuration**, click **OK**, and click **Search**.

Step 6 In the search result, check the historical configurations of Hive- and ZooKeeper-related services in the **Service** column. **Table 9-120** lists some configurations that may affect the connection between Hive and ZooKeeper.

Table 9-120 Configurations related to connection between Hive and ZooKeeper

Service	Parameter	Description
Hive	HIVE_GC_OPTS	HiveServer memory configuration. If the configuration is abnormal, HiveServer may restart repeatedly. In this case, you need to check the health status of the instance processes.
	hive.zookeeper.quorum	IP address of the node accommodating ZooKeeper that is connected to Hive.
	hive.zookeeper.client.port	Port of the ZooKeeper client connected to Hive.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Service	Parameter	Description
	hive.zookeeper.session.timeout	Timeout interval of the session set up between Hive and ZooKeeper.
	hive.zookeeper.connection.timeout	Timeout interval for Hive to connect to ZooKeeper.
	hive.zookeeper.connection.max.retries	Maximum number of retries for Hive to connect to ZooKeeper.
ZooKeeper	clientPort	Port number of the ZooKeeper client.
	ssl.enabled	Whether to enable SSL connections of ZooKeeper.

Restart related instances.

Step 7 Log in to FusionInsight Manager. Choose **O&M** > **Alarm** > **Alarms**, click the drop-down list in the row that contains the alarm, and view the role and the IP address of the node for which the alarm is generated in **Location**.

Step 8 Choose **Cluster**, click the name of the desired cluster, and choose **Services** > **Hive** > **Instance**. On the page that is displayed, select the instance at the IP address for which the alarm is generated, click **More**, and select **Restart Instance**.

NOTICE

During Hive instance restart, the instance cannot provide services for external systems. SQL tasks that are being executed on the instance may fail.

Step 9 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **Hive** for the target cluster.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.147 ALM-16048 Tez or Spark Library Path Does Not Exist

Description

The system checks the Tez and Spark library paths every 180 seconds. This alarm is generated when the Tez or Spark library path does not exist.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16048	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The Hive on Tez and Hive on Spark functions are affected.

Possible Causes

The Tez or Spark library path is deleted from the HDFS.

Procedure

Check the Tez and Spark library paths.

Step 1 Log in to the node where the client is deployed as user **root**.

Step 2 Run the following command to check whether the **tezlib** or **sparklib** directory exists in the **hdfs://hacluster/user/{User name}/.Trash/Current/** director:

hdfs dfs -ls hdfs://hacluster/user/*<username>***/.Trash/Current/**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, the following information shows that **/user/hive/tezlib/8.1.0.1/** and **/user/hive/sparklib/8.1.0.1/** exist.

```
host01:/opt/client # hdfs dfs -ls hdfs://hacluster/user/test/.Trash/Current/
Found 1 items
drwx------   - test hadoop          0 2019-06-17 19:53 hdfs://hacluster/user/test/.Trash/Current/user
```

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Run the following command to restore **tezlib** and **sparklib**.

hdfs dfs -mv hdfs://hacluster/user/*<username>*/.Trash/Current/user/hive/tezlib/8.1.0.1/tez.tar.gz /user/hive/tezlib/8.1.0.1/tez.tar.gz

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 Collect related information in the **.Trash/Current/** directory on the client background.

Step 6 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.148 ALM-17003 Oozie Service Unavailable

Description

The system checks the Oozie service status in every 5 seconds. This alarm is generated when Oozie or a component on which Oozie depends cannot provide services properly.

This alarm is automatically cleared when the Oozie service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
17003	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Details	Additional alarm information

Impact on the System

Oozie cannot be used to schedule jobs.

Possible Causes

- The DBService service is abnormal or the data of Oozie stored in DBService is damaged.
- The HDFS service is abnormal or the data of Oozie stored in HDFS is damaged.
- The YARN service is abnormal.
- The NodeAgent process is abnormal.

Procedure

Query the Oozie service health status code.

Step 1 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Oozie**. Click **oozie** (any one is OK) on the **oozie WebUI** to go to the Oozie WebUI.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 2 Add **/servicehealth** to the URL in the address box of the browser and access again. The value of **statusCode** is the current Oozie service health status code.

For example, visit **https://10.10.0.117:20026/Oozie/oozie/130/oozie/servicehealth**. The result is as follows:

```
{"beans":[{"name":"serviceStatus","statusCode":0}]}
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the health status code cannot be displayed or the browser does not respond, the service may be unavailable due to Oozie process fault. See **Step 13** to rectify the fault.

Step 3 Perform the operations based on the error code. For details, see **Table 9-121**.

Table 9-121 Oozie service health status code

Status Code	Description	Error Cause	Solution
0	The service is running properly.	None	None
18002	The DBService service is abnormal.	Oozie fails to connect to DBService or the data stored in DBService is damaged.	See **Step 4**.
18003	The HDFS service is abnormal.	Oozie fails to connect to HDFS or the data stored in HDFS is damaged.	See **Step 7**.
18005	The MapReduce service is abnormal.	The Yarn service is abnormal.	See **Step 11**.

Check the DBService service.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services**, and check whether the DBService service is running properly.

- If yes, go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Resolve the problem of DBService based on the alarm help and check whether the Oozie alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Step 6 Log in to the Oozie database to check whether the data is complete.

1. Log in to the active DBService node as user **root**.

   On the FusionInsight Manager page, choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService > Instance** to view the IP address of the active DBservice node.

2. Run the following command to log in to the Oozie database:

   su - omm

   source ${BIGDATA_HOME}/FusionInsight_BASE_8.1.0.1/install/FusionInsight-dbservice-2.7.0/.dbservice_profile

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

gsql -U *Username* **-W** *Oozie database password* **-p 20051 -d** *Database name*

3. After the login is successful, enter **\d** to check whether there are 15 data tables.

   The Oozie service has 15 data tables by default. If these data tables are deleted or the table structure is modified, the Oozie service may be unavailable. Contact the O&M personnel to back up the data and perform restoration.

Check the HDFS service.

Step 7 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services**, and check whether the HDFS service is running properly.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 8 Resolve the problem of HDFS based on the alarm help and check whether the Oozie alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Step 9 Log in to HDFS to check whether the Oozie file directory structure is complete.

1. Download and install an HDFS client..
2. Log in to the client node as user **root** and run the following commands to check whether **/user/oozie/share** exists.

   If the cluster uses the security mode, perform security authentication.

   kinit admin

   hdfs dfs -ls /user/oozie/share

- If yes, go to **Step 18**.
- If no, go to **Step 10**.

Step 10 In the Oozie client installation directory, manually upload the share directory to **/user/oozie** in HDFS, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Check the Yarn and MapReduce service.

Step 11 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* > **Services**, and check whether the Yarn and MapReduce services are running properly.

- If yes, go to **Step 18**.
- If no, go to **Step 12**.

Step 12 Resolve the problem of Yarn and MapReduce based on the alarm help and check whether the Oozie alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Check the Oozie process.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Log in to each node of Oozie as user **root**.

On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services**, click **Oozie**, and click the **Instances** tab. Record the IP address of the service nodes.

Step 14 Run the **ps -ef | grep oozie** command to check whether the Oozie process exists.

- If yes, go to **Step 15**.
- If no, go to **Step 18**.

Step 15 Collect fault information in **prestartDetail.log**, **oozie.log**, and **catalina.out** in the Oozie log directory **/var/log/Bigdata/oozie**. If the alarm is not caused by manual misoperation, go to **Step 16**.

Check the NodeAgent process.

Step 16 Log in to each node of Oozie as user **root**. Run the **ps -ef | grep nodeagent** command to check whether the NodeAgent process exists.

- If yes, go to **Step 17**.
- If no, go to **Step 18**.

Step 17 Run the **kill -9** *The process ID of nodeagent* command, wait 10 minutes, and check whether alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Step 18 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.149 ALM-17004 Oozie Heap Memory Usage Exceeds the Threshold

Description

The system checks the heap memory usage of the Oozie service every 60 seconds. The alarm is generated when the heap memory usage of a Metadata instance exceeds the threshold (95% of the maximum memory). The alarm is cleared when the heap memory usage is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
17004	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The heap memory overflow may cause a service breakdown. After the service breaks down, the Oozie service cannot be used to schedule tasks.

Possible Causes

The heap memory of the Oozie instance is overused or the heap memory is inappropriately allocated.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **Oozie Heap Memory Usage Exceeds the Threshold** > **Location**. Check the IP address of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* > **Services** > **Oozie** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **Memory** > **Oozie Heap Memory Resource Percentage**. Click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the used heap memory of Oozie reaches the threshold (the default value is 95% of the maximum heap memory) specified for Oozie.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Oozie** > **Configurations** > **All Configurations**. Set Search **GC_OPTS** in the search box. Increase the value of **-Xmx** as required, and click **Save** > **OK**.

NOTE

Suggestions on GC parameter settings for Oozie:

You are advised to set **-Xms** and **-Xmx** to the same value to prevent adverse impact on performance when JVM dynamically adjusts the heap memory size.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service or instance restart, services are interrupted, but submitted jobs are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Oozie** in the required cluster from the **Service**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.150 ALM-17005 Oozie Non Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the non heap memory usage of Oozie every 30 seconds. This alarm is reported if the non heap memory usage of Oozie exceeds the threshold

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

(80%). This alarm is cleared if the non heap memory usage is lower than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
17005	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The service may break down and the Oozie service cannot be used to schedule tasks.

Possible Causes

The non-heap memory of the Oozie instance is overused or the non-heap memory is inappropriately allocated.

Handling Procedure

Check non-heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **Oozie Non Heap Memory Usage Exceeds the Threshold**. On the displayed page, check the location information of the alarm. Check the name of the instance host for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Oozie** and click the **Instance** tab. On the displayed page, select the role corresponding to the host name for which the alarm is generated and select **Customize** from the drop-down list in the upper right corner of the chart area.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Choose **Memory** and select **Oozie Non Heap Memory Resource Percentage**. Click **OK**.

Step 3 Check whether the non-heap memory used by Oozie reaches the threshold (80% of the maximum non-heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Oozie** and click the **Configurations** and then **All Configurations**. On the displayed page, search for the **GC_OPTS** parameter in the search box and check whether it contains **-XX: MaxMetaspaceSize**. If yes, increase the value of **-XX: MaxMetaspaceSize** based on the site requirements. If no, manually add **-XX: MaxMetaspaceSize** and set its value to 1/8 of the value of **-Xmx**. Click **Save**, and then click **OK**

NOTE

JDK1.8 does not support the **MaxPermSize** parameter.

Suggestions on GC parameter settings for Oozie:

Set the value of **-XX:MaxMetaspaceSize** to 1/8 of the value of **-Xmx**. For example, if **-Xmx** is set to 2 GB, **-XX:MaxMetaspaceSize** is set to 256 MB. If **-Xmx** is set to 4 GB, **-XX:MaxMetaspaceSize** is set to 512 MB.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service or instance restart, services are interrupted, but submitted jobs are not affected.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Oozie** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.151 ALM-17006 Oozie Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the Oozie service every 30 seconds. The alarm is generated when the direct memory usage of an Oozie instance exceeds the threshold (80% of the maximum memory). The alarm is cleared when the direct memory usage of Oozie is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
17006	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The direct memory overflow may cause a service breakdown. After the service breaks down, the Oozie service cannot be used to schedule tasks.

Possible Causes

The direct memory of the Oozie instance is overused or the direct memory is inappropriately allocated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check direct memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Oozie Direct Memory Usage Exceeds the Threshold** > **Location**. Check the IP address of the instance involved in this alarm.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Oozie** and click the **Instance** tab. On the displayed page, select the role corresponding to the host name for which the alarm is generated and select **Customize** from the drop-down list in the upper right corner of the chart area. Choose **Memory** and select **Oozie Direct Buffer Resource Percentage**. Click **OK**.

Step 3 Check whether the used direct memory of Oozie reaches the threshold (the default value is 80% of the maximum direct memory) specified for Oozie.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 **On the FusionInsight Manager portal, choose Cluster > *Name of the desired cluster* > Services > Oozie > Configurations. Click All Configurations. Search GC_OPTS in the search box. Increase the value of -XX:MaxDirectMemorySize as required, and click Save. Click OK.**

NOTE

Suggestions on GC parameter settings for Oozie:

You are advised to set the value of **-XX:MaxDirectMemorySize** to 1/4 of the value of **-Xmx**. For example, if **-Xmx** is set to 4 GB, **-XX:MaxDirectMemorySize** is set to 1024 MB. If **-Xmx** is set to 2 GB, **-XX:MaxDirectMemorySize** is set to 512 MB. It is recommended that the value of **-XX:MaxDirectMemorySize** be greater than or equal to 512 MB.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service or instance restart, services are interrupted, but submitted jobs are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Oozie** in the required cluster from the **Service** drop-down list.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.152 ALM-17007 Garbage Collection (GC) Time of the Oozie Process Exceeds the Threshold

Description

The system checks GC time of the Oozie process every 60 seconds. The alarm is generated when GC time of the Oozie process exceeds the threshold (default value: **12 seconds**). The alarm is cleared when GC time is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
17007	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Oozie scheduling task responds slowly until the service is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The heap memory of the Oozie instance is overused or the heap memory is inappropriately allocated.

Procedure

Check GC time.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Garbage Collection (GC) Time of the Oozie Process Exceeds the Threshold** > **Location**. Check the IP address of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Oozie** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **GC** > **Garbage Collection (GC) Time of Oozie**. Click **OK**.

Step 3 Check whether GC time of the Oozie process every second exceeds the threshold (default value: **12 seconds**).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Oozie** > **Configurations. Click All Configurations. Search GC_OPTS in the search box. Increase the value of -Xmx as required, and click Save. Click OK.**

NOTE

Suggestions on GC parameter settings for Oozie:

You are advised to set **-Xms** and **-Xmx** to the same value to prevent adverse impact on performance when JVM dynamically adjusts the heap memory size.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service or instance restart, services are interrupted, but submitted jobs are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Oozie** in the required cluster from the **Service** drop-down list.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.153 ALM-17008 Abnormal Connection Between Oozie and ZooKeeper

Alarm Description

In HA mode, Oozie depends on ZooKeeper. This alarm is generated when the connection between Oozie and ZooKeeper is abnormal for three consecutive times.

This alarm is cleared when the connection between Oozie and ZooKeeper becomes normal.

This section applies to MRS 3.3.0 and later versions.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
17008	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Running scheduling tasks are blocked and new scheduling tasks cannot be submitted. In HA mode, the Oozie service will restart if this alarm is reported.

Possible Causes

- The ZooKeeper service is abnormal.
- Oozie fails to connect to ZooKeeper.

Handling Procedure

Check the ZooKeeper service status.

Step 1 In the service list on FusionInsight Manager, check whether **Running Status** of ZooKeeper is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 In the alarm list, check whether **ALM-13000 ZooKeeper Service Unavailable** is reported.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by performing the operations provided for **ALM-13000 ZooKeeper Service Unavailable**.

Step 4 Wait for several minutes and check whether the alarm **Abnormal Connection Between Oozie and ZooKeeper** is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the connectivity between Oozie and ZooKeeper.

Step 5 Log in to FusionInsight Manager, choose **O&M** > **Log** > **Online Search**, select the Oozie service, and search for the keyword **[Oozie Alarm Enhancement] [ZooKeeper]** in the log. View the cause in the log, and rectify the fault. In the alarm list, check whether the alarm **Abnormal Connection Between Oozie and ZooKeeper** is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Select **Oozie** for **Service** and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.154 ALM-17009 Abnormal Connection Between Oozie and DBService

Alarm Description

Oozie depends on DBService. After a task is submitted, the system checks DBService connectivity. This alarm is generated when the service fails the check for 10 consecutive times.

This alarm is cleared when the connection between Oozie and DBService becomes normal.

This section applies to MRS 3.3.0 and later versions.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
17009	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Running scheduling tasks are blocked and new scheduling tasks cannot be submitted.

Possible Causes

- The DBService service is abnormal.
- Oozie fails to connect to DBService.

Handling Procedure

Check the DBService status.

Step 1 In the service list on FusionInsight Manager, check whether **Running Status** of DBService is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 In the alarm list, check whether **ALM-27001 DBService Service Unavailable** is reported.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by performing the operations provided for **ALM-27001 DBService Service Unavailable**.

Step 4 Wait for several minutes and check whether the alarm **Abnormal Connection Between Oozie and DBService** is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the connectivity between Oozie and DBService.

Step 5 Log in to FusionInsight Manager, choose **O&M** > **Log** > **Online Search**, select the Oozie service, and search for the keyword **[Oozie Alarm Enhancement][DB Service]** in the log. View the cause in the log, and rectify the fault. In the alarm list, check whether the alarm **Abnormal Connection Between Oozie and DBService** is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Select **Oozie** for **Service** and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.155 ALM-17010 Abnormal Connection Between Oozie and HDFS

Alarm Description

Oozie depends on HDFS. After a task is submitted, the system checks HDFS connectivity. This alarm is generated when the service fails the check for 3 consecutive times.

This alarm is cleared when the connection between Oozie and HDFS becomes normal.

This section applies to MRS 3.3.0 and later versions.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
17010	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Running scheduling tasks are blocked and new scheduling tasks cannot be submitted.

Possible Causes

The HDFS service restarts, there is a fault, or the network connectivity is abnormal.

Handling Procedure

Check the HDFS service status.

Step 1 In the service list on FusionInsight Manager, check whether **Running Status** of HDFS is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 In the alarm list, check whether the "ALM-14000 HDFS Service Unavailable" alarm is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by performing the operations provided for **ALM-14000 HDFS Service Unavailable**.

Step 4 Wait for several minutes and check whether the alarm **Abnormal Connection Between Oozie and HDFS** is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the connectivity between Oozie and HDFS.

Step 5 Log in to FusionInsight Manager, choose **O&M** > **Log** > **Online Search**, select the Oozie service, and search for the keyword **[Oozie Alarm Enhancement][HDFS]** in the log. View the cause in the log, and rectify the fault. In the alarm list, check whether the alarm **Abnormal Connection Between Oozie and HDFS** is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Select **Oozie** for **Service** and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.156 ALM-17011 Abnormal Connection Between Oozie and YARN

Alarm Description

Oozie depends on YARN. After a task is submitted, the system checks YARN connectivity. This alarm is generated when the service fails the check for 5 consecutive times.

This alarm is cleared when the connection between Oozie and YARN becomes normal.

This section applies to MRS 3.3.0 and later versions.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
17011	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Running scheduling tasks are blocked and new scheduling tasks cannot be submitted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The Yarn service is abnormal.
- The connection between Oozie and Yarn is abnormal.

Handling Procedure

Check the YARN service status.

Step 1 In the service list on FusionInsight Manager, check whether **Running Status** of Yarn is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 In the alarm list, check whether **ALM-18000 YARN Service Unavailable** is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by performing the operations provided for **ALM-18000 YARN Service Unavailable**.

Step 4 Wait for several minutes and check whether the alarm **Abnormal Connection Between Oozie and Yarn** is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the connectivity between Oozie and Yarn.

Step 5 Log in to FusionInsight Manager, choose **O&M** > **Log** > **Online Search**, select the Oozie service, and search for the keyword **[Oozie Alarm Enhancement][Yarn]** in the log. View the cause in the log, and rectify the fault. In the alarm list, check whether the alarm **Abnormal Connection Between Oozie and Yarn** is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Select **Oozie** for **Service** and click **OK**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None.

9.11.157 ALM-18000 Yarn Service Unavailable

Alarm Description

The alarm module checks the Yarn service status every 60 seconds. This alarm is generated when the Yarn service is unavailable.

The alarm is cleared when the Yarn service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18000	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The cluster cannot provide the Yarn service.
- Users cannot run new applications.
- Submitted applications cannot be run.

Possible Causes

- The ZooKeeper service is abnormal.
- The HDFS service is abnormal.
- There is no active ResourceManager instance in the Yarn cluster.
- All the NodeManager nodes in the Yarn cluster are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check ZooKeeper service status.

Step 1 On the FusionInsight Manager, check whether the alarm list contains **ALM-13000 ZooKeeper Service Unavailable**.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Rectify the fault by performing the operations provided in **ALM-13000 ZooKeeper Service Unavailable**. Then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check the HDFS service status.

Step 3 On the FusionInsight Manager, check whether the alarm list contains the HDFS alarms.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Choose **O&M > Alarm > Alarms**, handle HDFS alarms based on the alarm help, and check whether the Yarn alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the ResourceManager status in the Yarn cluster.

Step 5 On FusionInsight Manager, choose **Cluster** > **Services** > **Yarn**.

Step 6 On the **Dashboard** page, check whether there is any active ResourceManager instance in the Yarn cluster.

- If yes, go to **Step 7**.
- If no, go to **Step 10**.

Check the NodeManager node status in the Yarn cluster.

Step 7 On FusionInsight Manager, choose **Cluster** > **Services** > **Yarn** > **Instances**.

Step 8 Query the status of NodeManager and check whether there are unhealthy nodes.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Rectify the fault by following the steps provided in **ALM-18002 NodeManager Heartbeat Lost** or **ALM-18003 NodeManager Unhealthy**. After the fault is rectified, check whether the Yarn alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On the FusionInsight Manager portal of the active cluster, choose **O&M** > **Log > Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 Select **Yarn** in the required cluster from the **Service**.

Step 12 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 13 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.158 ALM-18002 NodeManager Heartbeat Lost

Alarm Description

The system checks the number of lost NodeManager nodes every 30 seconds, and compares the number with the threshold. The **Lost Nodes** indicator has a default threshold. The alarm is generated when the value of Number of Lost Nodes exceeds the threshold.

To change the threshold, on FusionInsight Manager, choose **Cluster** > **Services** > **Yarn**. On the displayed page, choose **Configurations** > **All Configurations**, and change the value of **yarn.nodemanager.lost.alarm.threshold**. You do not need to restart Yarn to make the change take effect.

The default threshold is 0. The alarm is generated when the number of lost nodes exceeds the threshold, and is cleared when the number of lost nodes is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18002	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Lost Host	Specifies the list of hosts with lost nodes.

Impact on the System

- The lost NodeManager node cannot provide the Yarn service.
- The number of containers decreases, so the cluster performance deteriorates.

Possible Causes

- NodeManager is forcibly deleted without decommission.
- All the NodeManager instances are stopped or the NodeManager process is faulty.
- The host where the NodeManager node resides is faulty.
- The network between the NodeManager and ResourceManager is disconnected or busy.

Handling Procedure

Check the NodeManager status.

Step 1 On the FusionInsight Manager, and choose **O&M** > **Alarm > Alarms**. Click before the alarm and obtain lost nodes in **Additional Information**.

Step 2 Check whether the lost nodes are hosts that have been manually deleted without decommission.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 After the setting, Choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn**. On the displayed page, choose **Configurations** > **All Configurations**. Search for **yarn.nodemanager.lost.alarm.threshold** and change its value to the number of hosts that are not out of service and proactively deleted. After the setting, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Step 4 Manually clear the alarm. Note that decommission must be performed before deleting hosts.

Step 5 On the FusionInsight Manager portal, choose **Cluster > Hosts**, and check whether the nodes obtained in **Step 1** are healthy.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

Step 6 Rectify the node fault based on **ALM-12006 Node Fault** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check the process status.

Step 7 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **Instance**, and check whether there are NodeManager instances whose status is not **Good**.

- If yes, go to **Step 10**.
- If no, go to **Step 8**.

Step 8 Check whether the NodeManager instance is deleted.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Restart the active and standby ResourceManager instances, and check whether the alarm is cleared.

NOTE

- Restarting the active ResourceManager instance will trigger a active/standby switchover of the ResourceManager instance. During the switchover, Yarn cannot submit new jobs, but submitted jobs are not affected. The Yarn component and components that depend on Yarn generate temporary service unavailability alarms.
- Restart the standby ResourceManager instance. Services are not affected.

- If yes, no further action is required.
- If no, go to **Step 13**.

Check the instance status.

Step 10 Select NodeManager instances which running state is not **Normal** and restart them. Check whether the alarm is cleared.

NOTE

During NodeManager restart, containers submitted to this node may be retried to other nodes.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check the network status.

Step 11 Log in to the management node, **ping** the IP address of the lost NodeManager node to check whether the network is disconnected or busy.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 Rectify the network, and check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 13**.

Collect fault information.

Step 13 On the FusionInsight Manager in the active cluster, choose **O&M** > **Log** > **Download**.

Step 14 Select **Yarn** in the required cluster from the **Service**.

Step 15 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 16 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.159 ALM-18003 NodeManager Unhealthy

Description

The system checks the number of unhealthy NodeManager nodes every 30 seconds, and compares the number with the threshold. The Unhealthy Nodes indicator has a default threshold. This alarm is generated when the value of the Unhealthy Nodes indicator exceeds the threshold.

To change the threshold, on FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn**. On the displayed page, choose **Configurations** > **All Configurations**, and change the value

of **yarn.nodemanager.unhealthy.alarm.threshold**. You do not need to restart Yarn to make the change take effect.

The default threshold is 0. The alarm is generated when the number of unhealthy nodes exceeds the threshold, and is cleared when the number of unhealthy nodes is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18003	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Unhealthy Host	Specifies the list of hosts with unhealthy nodes.

Impact on the System

- The faulty NodeManager node cannot provide the Yarn service.
- The number of containers decreases, so the cluster performance deteriorates.

Possible Causes

- The hard disk space of the host where the NodeManager node resides is insufficient.
- User **omm** does not have the permission to access a local directory on the NodeManager node.

Procedure

Check the hard disk space of the host.

Step 1 On the FusionInsight Manager, and choose **O&M** > **Alarm > Alarms**. Click before the alarm and obtain unhealthy nodes in **Additional Information**.

Step 2 Choose **Cluster** > *Name of the desired cluster* **> Services > Yarn** > **Instance**, select the NodeManager instance corresponding to the host, choose **Instance Configurations > All Configurations** and view disks corresponding to **yarn.nodemanager.local-dirs** and **yarn.nodemanager.log-dirs**.

Step 3 Choose **O&M > Alarm > Alarms**. In the alarm list, check whether the related disk has the alarm **ALM-12017 Insufficient Disk Capacity**.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Rectify the fault by following the instructions provided in **ALM-12017 Insufficient Disk Capacity**. After the fault is rectified, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Choose **Hosts** > *Name of the desired host* . On the **Dashboard** page, check the disk usage of the corresponding partition. Check whether the percentage of the used space of the mounted disk exceeds the value of **yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-percentage**

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 Reduce the disk usage to less than the configured value, wait for 10 to 20 minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check the access permission of the local directory on each NodeManager node.

Step 7 Obtain the NodeManager directory viewed in **Step 2**, log in to each NodeManager node as user **root**, and go to the obtained directory.

Step 8 Run the **ll** command to check whether the permission of the **localdir** and **containerlogs** folders is **755** and whether **User:Group** is **omm:ficommon**.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Run the following command to set the permission to **755** and **User:Group** to **omm:ficommon**:

chmod 755 *<folder_name>*

chown omm:ficommon *<folder_name>*

Step 10 Wait for 10 to 20 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager in the active cluster, choose **O&M** > **Log** > **Download**.

Step 12 Select **Yarn** in the required cluster from the **Service**.

Step 13 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.160 ALM-18008 Heap Memory Usage of ResourceManager Exceeds the Threshold

Description

The system checks the heap memory usage of Yarn ResourceManager every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the heap memory usage of Yarn ResourceManager exceeds the threshold (95% of the maximum memory by default).

Users can choose **O&M > Alarm > Thresholds** > *Name of the desired cluster* > **Yarn** to change the threshold.

When the **Trigger Count** is 1, this alarm is cleared when the heap memory usage of Yarn ResourceManager is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the heap memory usage of Yarn ResourceManager is less than or equal to 95% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18008	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the heap memory usage of Yarn ResourceManager is overhigh, the performance of Yarn task submission and operation is affected. In addition, a memory overflow may occur so that the Yarn service is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The heap memory of the Yarn ResourceManager instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **Heap Memory Usage of Yarn ResourceManager Exceeds the Threshold** > **Location**. Check the HostName of the instance for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* **> Services** > **Yarn** > **Instance** > **ResourceManager** (Indicates the host name of the instance for which the alarm is generated)**.** Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **ResourceManager > Percentage of Used Memory of the ResourceManager**. Check the heap memory usage.

Step 3 Check whether the used heap memory of ResourceManager reaches 95% of the maximum heap memory specified for ResourceManager.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* **> Services** > **Yarn** > **Configurations** > **All Configurations** > **ResourceManager** > **System**. Increase the value of **GC_OPTS** parameter as required, click **Save**. Restart the role instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- Restarting the active ResourceManager instance will trigger a active/standby switchover of the ResourceManager instance. During the switchover, Yarn cannot submit new jobs, but submitted jobs are not affected. The Yarn component and components that depend on Yarn generate temporary service unavailability alarms.

  Restart the standby ResourceManager instance. Services are not affected.
- The mapping between the number of NodeManager instances in a cluster and the memory size of ResourceManager is as follows:
  - If the number of NodeManager instances in the cluster reaches 100, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 200, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 500, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms10G -Xmx10G -XX:NewSize=1G -XX:MaxNewSize=2G
  - If the number of NodeManager instances in the cluster reaches 1000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms20G -Xmx20G -XX:NewSize=1G -XX:MaxNewSize=2G
  - If the number of NodeManager instances in the cluster reaches 2000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms40G -Xmx40G -XX:NewSize=2G -XX:MaxNewSize=4G
  - If the number of NodeManager instances in the cluster reaches 3000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms60G -Xmx60G -XX:NewSize=2G -XX:MaxNewSize=4G
  - If the number of NodeManager instances in the cluster reaches 4000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms80G -Xmx80G -XX:NewSize=2G -XX:MaxNewSize=4G
  - If the number of NodeManager instances in the cluster reaches 5000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms100G -Xmx100G -XX:NewSize=3G -XX:MaxNewSize=6G

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select the following node in the required cluster from the **Service**.

- NodeAgent
- Yarn

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.161 ALM-18009 Heap Memory Usage of JobHistoryServer Exceeds the Threshold

Description

The system checks the heap memory usage of MapReduce JobHistoryServer every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the heap memory usage of MapReduce JobHistoryServer exceeds the threshold (95% of the maximum memory by default).

Users can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **MapReduce** to change the threshold.

When the **Trigger Count** is 1, this alarm is cleared when the heap memory usage of MapReduce JobHistoryServer is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the heap memory usage of MapReduce JobHistoryServer is less than or equal to 95% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18009	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the heap memory usage of MapReduce JobHistoryServer is overhigh, the performance of MapReduce log archiving is affected. What is more, a memory overflow occurs so that the YARN service is unavailable.

Possible Causes

The heap memory of the MapReduce JobHistoryServer instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **ALM-18009 Heap Memory Usage of MapReduce JobHistoryServer Exceeds the Threshold** > **Location**. Check the HostName of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **MapReduce** and click the **Instance** tab. On the displayed page, select **JobHistoryServer** (the host name for which the alarm is generated) and select **Customize** > **Resource** from the drop-down list in the upper right corner of the chart area. On the displayed page, select **JobHistoryServer Heap Memory Usage Statistics**. Check the heap memory usage.

Step 3 Check whether the used heap memory of JobHistoryServer reaches 95% of the maximum heap memory specified for JobHistoryServer.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* **> Services** > **MapReduce** > **Configurations** > **All Configurations** > **JobHistoryServer** > **System**. Increase the value of **GC_OPTS** parameter as required, click **Save**. Click **OK** and restart the role instance.

NOTE

- The mapping between the number of historical tasks (10000) and the memory of JobHistoryServer is as follows:

  -Xms30G -Xmx30G -XX:NewSize=1G -XX:MaxNewSize=2G

- During the restart of JobHistoryServer, the status query of tasks such as Hive is affected, and the query result may be inaccurate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select the following node in the required cluster from the **Service**.

- NodeAgent
- MapReduce

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.162 ALM-18010 ResourceManager GC Time Exceeds the Threshold

Description

The system checks the garbage collection (GC) duration of the ResourceManager process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default).

This alarm is cleared when the GC duration is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18010	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A long GC duration of the ResourceManager process may interrupt the services.

Possible Causes

The heap memory of the ResourceManager instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-18010 ResourceManager GC Time Exceeds the Threshold** > **Location**. View the IP address of the alarmed instance.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Yarn**. On the page that is displayed, click the **Instance** tab. In the instance list, select **ResourceManager** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart, choose **Customize** > **Garbage Collection**, and select **Garbage Collection (GC) Time of ResourceManager**. Check the GC duration statistics of the ResourceManager process every minute.

Step 3 Check whether the GC duration of the ResourceManager process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **Yarn** > **Configurations** > **All Configurations** > **ResourceManager** > **System**. Increase the value of the **GC_OPTS** parameter as required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The mapping between the number of NodeManager instances in a cluster and the memory size of ResourceManager is as follows:

- If the number of NodeManager instances in the cluster reaches 100, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
- If the number of NodeManager instances in the cluster reaches 200, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=1G
- If the number of NodeManager instances in the cluster reaches 500, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms10G -Xmx10G -XX:NewSize=1G -XX:MaxNewSize=2G
- If the number of NodeManager instances in the cluster reaches 1000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms20G -Xmx20G -XX:NewSize=1G -XX:MaxNewSize=2G
- If the number of NodeManager instances in the cluster reaches 2000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms40G -Xmx40G -XX:NewSize=2G -XX:MaxNewSize=4G
- If the number of NodeManager instances in the cluster reaches 3000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms60G -Xmx60G -XX:NewSize=2G -XX:MaxNewSize=4G
- If the number of NodeManager instances in the cluster reaches 4000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms80G -Xmx80G -XX:NewSize=2G -XX:MaxNewSize=4G
- If the number of NodeManager instances in the cluster reaches 5000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms100G -Xmx100G -XX:NewSize=3G -XX:MaxNewSize=6G

Step 5 Save the configuration and restart the ResourceManager instance.

NOTE

- Restarting the active ResourceManager instance will trigger a active/standby switchover of the ResourceManager instance. During the switchover, Yarn cannot submit new jobs, but submitted jobs are not affected. The Yarn component and components that depend on Yarn generate service unavailable alarms for a short period of time.
- Restart the standby ResourceManager instance. Services are not affected.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **ResourceManager** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.163 ALM-18011 NodeManager GC Time Exceeds the Threshold

Description

The system checks the garbage collection (GC) duration of the NodeManager process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default).

This alarm is cleared when the GC duration is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18011	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A long GC duration of the NodeManager process may interrupt the services.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The heap memory of the NodeManager instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-18011 NodeManager GC Time Exceeds the Threshold** > **Location**. View the IP address of the alarmed instance.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Yarn**. On the page that is displayed, click the **Instance** tab. In the instance list, select **NodeManager** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart, choose **Customize** > **Garbage Collection**, and select **Garbage Collection (GC) Time of NodeManager**. Check the GC duration statistics of the NodeManager process every minute.

Step 3 Check whether the GC duration of the NodeManager process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **Yarn** > **Configurations** > **All Configurations** > **NodeManager** > **System**. Increase the value of the **GC_OPTS** parameter as required.

NOTE

The mapping between the number of NodeManager instances in a cluster and the memory size of NodeManager is as follows:

- If the number of NodeManager instances in the cluster reaches 100, the recommended JVM parameters for NodeManager instances are as follows: -Xms2G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
- If the number of NodeManager instances in the cluster reaches 200, the recommended JVM parameters for NodeManager instances are as follows: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
- If the number of NodeManager instances in the cluster reaches 500, the recommended JVM parameters for NodeManager instances are as follows: -Xms8G -Xmx8G -XX:NewSize=1G -XX:MaxNewSize=2G

Step 5 Save the configuration and restart the NodeManager instance.

NOTE

During NodeManager restart, containers submitted to this node may be retried to other nodes.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **NodeManager** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.164 ALM-18012 JobHistoryServer GC Time Exceeds the Threshold

Description

The system checks the garbage collection (GC) duration of the JobHistoryServer process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default).

This alarm is cleared when the GC duration is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18012	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A long GC duration of the JobHistoryServer process may interrupt the services.

Possible Causes

The heap memory of the JobHistoryServer instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-18012 JobHistoryServer GC Time Exceeds the Threshold** > **Location**. View the IP address of the alarmed instance.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **MapReduce**. On the page that is displayed, click the **Instance** tab. In the instance list, select **JobHistoryServer** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart, choose **Customize** > **Garbage Collection**, and select **Garbage Collection (GC) Time of the JobHistoryServer**. Check the GC duration statistics of the JobHistoryServer process every minute.

Step 3 Check whether the GC duration of the JobHistoryServer process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* **> Services** > **MapReduce** > **Configurations** > **All Configurations** > **JobHistoryServer** > **System**. Increase the value of the **GC_OPTS** parameter as required.

NOTE

The mapping between the number of historical tasks (10000) and the memory of the JobHistoryServer is as follows:

-Xms30G -Xmx30G -XX:NewSize=1G -XX:MaxNewSize=2G

Step 5 Save the configuration and restart the JobHistoryServer instance.

NOTE

During the restart of JobHistoryServer, the status query of tasks such as Hive is affected, and the query result may be inaccurate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **JobHistoryServer** in the required cluster from the **Service**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.165 ALM-18013 ResourceManager Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of ResourceManager every 30 seconds. This alarm is generated when the direct memory usage of ResourceManager instances exceeds the threshold (90% of the maximum memory).

This alarm is automatically cleared when the direct memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18013	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the available direct memory of ResourceManager is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The direct memory of ResourceManager instances is overused or the direct memory is inappropriately allocated.

Handling Procedure

Check the direct memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-18013 ResourceManager Direct Memory Usage Exceeds the Threshold** > **Location**. View the IP address of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **Yarn**. On the page that is displayed, click the **Instances** tab and click the ResourceManager instance for which this alarm is generated. Click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Resource**, and select **Memory Usage Status of ResourceManager** to check the direct memory usage.

Step 3 Check whether the used direct memory of a ResourceManager instance reaches 90% (default threshold) of the maximum direct memory allocated to it.

- If yes, go to **Step 4**.
- If no, go to **Step 9**.

Step 4 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **Yarn** > **Configurations** > **All Configurations** > **ResourceManager** > **System**. Check whether **-XX:MaxDirectMemorySize** exists in the **GC_OPTS** parameter.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Delete the **-XX:MaxDirectMemorySize** parameter from **GC_OPTS** and save the configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

MaxDirectMemorySize indicates the maximum off-heap memory size. If the **MaxDirectMemorySize** parameter of ResourceManager is not specified, the memory of ResourceManager is not limited. By default, **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter is not set.

Step 6 Perform the following steps to restart the ResourceManager instance:

NOTICE

- Restarting the standby ResourceManager instance does not affect services.
- During the ResourceManager switchover, new jobs cannot be submitted to Yarn, but submitted jobs are not affected.

1. On the Yarn service page, click the **Instances** tab, select the **ResourceManager (Standby)** instance, choose **More**, select **Restart Instance**, and verify the password to restart the instance.
2. After the standby instance is restarted, click the **Dashboard** tab of Yarn, choose **More**, select **Perform ResourceManager Switchover**, and verify the password to perform an active/standby switchover.
3. After the active/standby switchover is complete, click the **Instances** tab on the Yarn service page, select the **ResourceManager (Standby)** instance, choose **More**, select **Restart Instance**, and verify the password to restart the instance. Wait until the instance is restarted.

Step 7 Check whether **ALM-18008 Heap Memory Usage of ResourceManager Exceeds the Threshold** exists.

- If yes, rectify the fault by referring to **ALM-18008 Heap Memory Usage of ResourceManager Exceeds the Threshold**.
- If no, go to **Step 8**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **ResourceManager** for the target cluster.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.166 ALM-18014 NodeManager Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the Yarn service every 30 seconds. This alarm is generated when the direct memory usage of a NodeManager instance exceeds the threshold (90% of the maximum memory).

The alarm is cleared when the direct memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18014	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available direct memory of the Yarn service is insufficient, a memory overflow occurs and the service breaks down.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The direct memory of the NodeManager instance is overused or the direct memory is inappropriately allocated.

Procedure

Check the direct memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-18014 NodeManager Direct Memory Usage Exceeds the Threshold** > **Location**. View the IP address of the alarmed instance.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn**. On the page that is displayed, click the **Instance** tab. In the instance list, select **NodeManager** (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart, choose **Customize** > **Resource**, and select **Percentage of Used Memory of the NodeManager**. Check the direct memory usage.

Step 3 Check whether the used direct memory of NodeManager reaches 90% of the maximum direct memory specified for NodeManager by default.

- If yes, go to **Step 4**.
- If no, go to **Step 9**.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **Yarn** > **Configurations** > **All Configurations** > **NodeManager** > **System**. Check whether **-XX:MaxDirectMemorySize** exists in the **GC_OPTS** parameter.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 In the **GC_OPTS** parameter, delete "-XX:MaxDirectMemorySize".

Step 6 Save the configuration and restart the NodeManager instance.

NOTE

During NodeManager restart, containers submitted to this node may be retried to other nodes.

Step 7 Check whether the **ALM-18018 NodeManager Heap Memory Usage Exceeds the Threshold** exists.

- If yes, handle the alarm by referring to **ALM-18018 NodeManager Heap Memory Usage Exceeds the Threshold**.
- If no, go to **Step 8**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Select **NodeManager** in the required cluster from the **Service**.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.167 ALM-18015 JobHistoryServer Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the MapReduce service every 30 seconds. This alarm is generated when the direct memory usage of a JobHistoryServer instance exceeds the threshold (90% of the maximum memory).

The alarm is cleared when the direct memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18015	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available direct memory of the MapReduce service is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The direct memory of the JobHistoryServer instance is overused or the direct memory is inappropriately allocated.

Procedure

Check the direct memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-18015 JobHistory Direct Memory Usage Exceeds the Threshold** > **Location**. View the IP address of the alarmed instance.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **MapReduce** and click the **Instance** tab. On the displayed page, select **JobHistoryServer** (the IP address of the instance for which the alarm is generated). Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **Resource**, select **Memory Usage Status of JobHistoryServer**. Check the direct memory usage.

Step 3 Check whether the used direct memory of MapReduce reaches 90% of the maximum direct memory specified for the MapReduce service.

- If yes, go to **Step 4**.
- If no, go to **Step 9**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* **> Services** > **MapReduce** > **Configurations** > **All Configurations** > **JobHistoryServer** > **System**. Check whether **-XX:MaxDirectMemorySize** exists in the **GC_OPTS** parameter.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 In the **GC_OPTS** parameter, delete "-XX:MaxDirectMemorySize".

Step 6 Save the configuration and restart the JobHistoryServer instance.

NOTE

During the restart of JobHistoryServer, the status query of tasks such as Hive is affected, and the query result may be inaccurate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Check whether the **ALM-18009 Heap Memory Usage of JobHistoryServer Exceeds the Threshold** exists.

- If yes, handle the alarm by referring to **ALM-18009 Heap Memory Usage of JobHistoryServer Exceeds the Threshold**.
- If no, go to **Step 8**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 10 Select **JobHistoryServer** in the required cluster from the **Service**.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.168 ALM-18016 Non Heap Memory Usage of ResourceManager Exceeds the Threshold

Description

The system checks the Non Heap memory usage of Yarn ResourceManager every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the Non Heap memory usage of Yarn ResourceManager exceeds the threshold (90% of the maximum memory by default).

Users can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Yarn** to change the threshold.

The alarm is cleared when the Non Heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18016	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the Non Heap memory usage of Yarn ResourceManager is overhigh, the performance of Yarn task submission and operation is affected. In addition, a memory overflow may occur so that the Yarn service is unavailable.

Possible Causes

The Non Heap memory of the Yarn ResourceManager instance on the node is overused or the Non Heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the Non Heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **ALM-18016 Non Heap Memory Usage of Yarn ResourceManager Exceeds the Threshold** > **Location**. Check the HostName of the instance for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **Yarn** > **Instance** > **ResourceManager**. Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **Percentage of Used Memory of the ResourceManager**. ResourceManager indicates the corresponding HostName of the instance for which the alarm is generated. Check the Non Heap memory usage.

Step 3 Check whether the used Non Heap memory of ResourceManager reaches 90% of the maximum Non Heap memory specified for ResourceManager by default.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **Configurations** > **All Configurations** > **ResourceManager** > **System**. Adjust the **GC_OPTS** memory parameter of ResourceManager. Save the configuration and restart the ResourceManager instance.

NOTE

- Restarting the active ResourceManager instance will trigger a active/standby switchover of the ResourceManager instance. During the switchover, Yarn cannot submit new jobs, but submitted jobs are not affected. The Yarn component and components that depend on Yarn generate service unavailable alarms for a short period of time.

  Restart the standby ResourceManager instance. Services are not affected.
- The mapping between the number of NodeManager instances in a cluster and the memory size of ResourceManager is as follows:
  - If the number of NodeManager instances in the cluster reaches 100, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 200, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 500, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms10G -Xmx10G -XX:NewSize=1G -XX:MaxNewSize=2G
  - If the number of NodeManager instances in the cluster reaches 1000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms20G -Xmx20G -XX:NewSize=1G -XX:MaxNewSize=2G
  - If the number of NodeManager instances in the cluster reaches 2000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms40G -Xmx40G -XX:NewSize=2G -XX:MaxNewSize=4G
  - If the number of NodeManager instances in the cluster reaches 3000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms60G -Xmx60G -XX:NewSize=2G -XX:MaxNewSize=4G
  - If the number of NodeManager instances in the cluster reaches 4000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms80G -Xmx80G -XX:NewSize=2G -XX:MaxNewSize=4G
  - If the number of NodeManager instances in the cluster reaches 5000, the recommended JVM parameters of the ResourceManager instance are as follows: -Xms100G -Xmx100G -XX:NewSize=3G -XX:MaxNewSize=6G

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select the following node in the required cluster from the **Service**.

- NodeAgent
- Yarn

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.169 ALM-18017 Non Heap Memory Usage of NodeManager Exceeds the Threshold

Description

The system checks the Non Heap memory usage of Yarn NodeManager every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the Non Heap memory usage of Yarn NodeManager exceeds the threshold (90% of the maximum memory by default).

Users can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Yarn** to change the threshold.

The alarm is cleared when the Non Heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18017	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the Non Heap memory usage of Yarn NodeManager is overhigh, the performance of Yarn task submission and operation is affected. In addition, a memory overflow may occur so that the Yarn service is unavailable.

Possible Causes

The Non Heap memory of the Yarn NodeManager instance on the node is overused or the Non Heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the Non Heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **ALM-18017 Non Heap Memory Usage of Yarn NodeManager Exceeds the Threshold** > **Location**. Check the HostName of the instance for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **Yarn** > **Instance** > **NodeManager**. Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **Resource** > **Percentage of Used Memory of the NodeManager**. NodeManager indicates the corresponding HostName of the instance for which the alarm is generated. Check the Non Heap memory usage.

Step 3 Check whether the used Non Heap memory of NodeManager reaches 90% of the maximum Non Heap memory specified for NodeManager by default.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **Yarn** > **Configurations** > **All Configurations** > **NodeManager** > **System**. Adjust the **GC_OPTS** memory parameter of NodeManager, click **Save**, and click **OK,** and restart the role instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- During NodeManager restart, containers submitted to this node may be retried to other nodes.
- The mapping between the number of NodeManager instances in a cluster and the memory size of NodeManager is as follows:
  - If the number of NodeManager instances in the cluster reaches 100, the recommended JVM parameters for NodeManager instances are as follows: -Xms2G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 200, the recommended JVM parameters for NodeManager instances are as follows: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 500, the recommended JVM parameters for NodeManager instances are as follows: -Xms8G -Xmx8G -XX:NewSize=1G -XX:MaxNewSize=2G

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select the following node in the required cluster from the **Service**.

- NodeAgent
- Yarn

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.170 ALM-18018 NodeManager Heap Memory Usage Exceeds the Threshold

Description

The system checks the heap memory usage of Yarn NodeManager every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the heap memory usage of Yarn NodeManager exceeds the threshold (95% of the maximum memory by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The alarm is cleared when the heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18018	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the heap memory usage of Yarn NodeManager is overhigh, the performance of Yarn task submission and operation is affected. In addition, a memory overflow may occur so that the Yarn service is unavailable.

Possible Causes

The heap memory of the Yarn NodeManager instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **ALM-18018 NodeManager Heap Memory Usage Exceeds the Threshold** > **Location**. Check the HostName of the instance for which the alarm is generated.

Step 2 On the **Home** page of FusionInsight Manager, choose **Cluster** > **Services** > **Yarn**. On the page that is displayed, click the **Instance** tab. In the instance list, select

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NodeManager (IP address of the instance for which this alarm is generated). Click the drop-down list in the upper right corner of the chart, choose **Customize** > **Resource**, and select **Percentage of Used Memory of the NodeManager**. Check the heap memory usage.

Step 3 Check whether the used heap memory of NodeManager reaches 95% of the maximum heap memory specified for NodeManager.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **Yarn** > **Configurations** > **All Configurations** > **NodeManager** > **System**. Increase the value of the **GC_OPTS** parameter as required. Save the configurations and restart the NodeManager instance.

NOTE

- During NodeManager restart, containers submitted to this node may be retried to other nodes.
- The mapping between the number of NodeManager instances in a cluster and the memory size of NodeManager is as follows:
  - If the number of NodeManager instances in the cluster reaches 100, the recommended JVM parameters for NodeManager instances are as follows: -Xms2G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 200, the recommended JVM parameters for NodeManager instances are as follows: -Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G
  - If the number of NodeManager instances in the cluster reaches 500, the recommended JVM parameters for NodeManager instances are as follows: -Xms8G -Xmx8G -XX:NewSize=1G -XX:MaxNewSize=2G

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select the following node in the required cluster from the **Service**.

- NodeAgent
- Yarn

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.171 ALM-18019 Non Heap Memory Usage of JobHistoryServer Exceeds the Threshold

Description

The system checks the Non Heap memory usage of MapReduce JobHistoryServer every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the Non Heap memory usage of MapReduce JobHistoryServer exceeds the threshold (90% of the maximum memory by default).

Users can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **MapReduce** to change the threshold.

The alarm is cleared when the Non Heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18019	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the Non Heap memory usage of MapReduce JobHistoryServer is overhigh, the performance of MapReduce task submission and operation is affected. In

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

addition, a memory overflow may occur so that the MapReduce service is unavailable.

Possible Causes

The Non Heap memory of the MapReduce JobHistoryServer instance on the node is overused or the Non Heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the Non Heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** > **ALM-18019 Non Heap Memory Usage of MapReduce JobHistoryServer Exceeds the Threshold** > **Location**. Check the HostName of the instance for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **MapReduce** > **Instance** > **JobHistoryServer.** Click the drop-down menu in the upper right corner of **Chart**, choose **Customize** > **JobHistoryServer Non Heap memory usage statistics**. JobHistoryServer indicates the corresponding HostName of the instance for which the alarm is generated. Check the Non Heap memory usage.

Step 3 Check whether the used Non Heap memory of JobHistoryServer reaches 90% of the maximum Non Heap memory specified for JobHistoryServer.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster >** *Name of the desired cluster* **> Services** > **MapReduce** > **Configurations** > **All Configurations** > **JobHistoryServer** > **System**. Adjust the **GC_OPTS** memory parameter of the NodeManager, click **Save**, and click **OK,** and restart the role instance.

NOTE

- During the restart of JobHistoryServer, the status query of tasks such as Hive is affected, and the query result may be inaccurate.
- The mapping between the number of historical tasks (10000) and the memory of the JobHistoryServer is as follows:

  -Xms30G -Xmx30G -XX:NewSize=1G -XX:MaxNewSize=2G

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select the following node in the required cluster from the **Service**.

- NodeAgent
- MapReduce

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.172 ALM-18020 Yarn Task Execution Timeout

Alarm Description

The system checks MapReduce and Spark tasks (except for permanent JDBC tasks) submitted to Yarn every 15 minutes. This alarm is generated when the task execution time exceeds the timeout duration specified by the user. However, the task can be properly executed. The client timeout parameter of MapReduce is mapreduce.application.timeout.alarm and that of Spark is spark.application.timeout.alarm. The unit is ms.

This alarm is cleared when the task is finished or terminated.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18020	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
ApplicationName	Specifies the object (application ID) for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The alarm persists after task execution times out. However, the task can still be properly executed, so this alarm does not exert any impact on the system.

Possible Causes

- The specified timeout duration is shorter than the required execution time.
- The queue resources for task running are insufficient.
- Task data skew occurs. As a result, some tasks process a large amount of data and take a long time to execute.

Handling Procedure

Check whether the timeout interval is correctly set.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. The **Alarms** page is displayed.

Step 2 Select the alarm whose ID is **18020**. In the alarm details, view **Location** to obtain the timeout task name and timeout duration.

Step 3 Based on the task name and timeout interval, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **ResourceManager (Active)** to log in to the native Yarn page. Then find the task on the native page, check its **StartTime** and calculate the task execution time based on the current system time. Check whether the task execution time exceeds the timeout duration.

- If yes, go to **Step 4**.
- If no, go to **Step 10**.

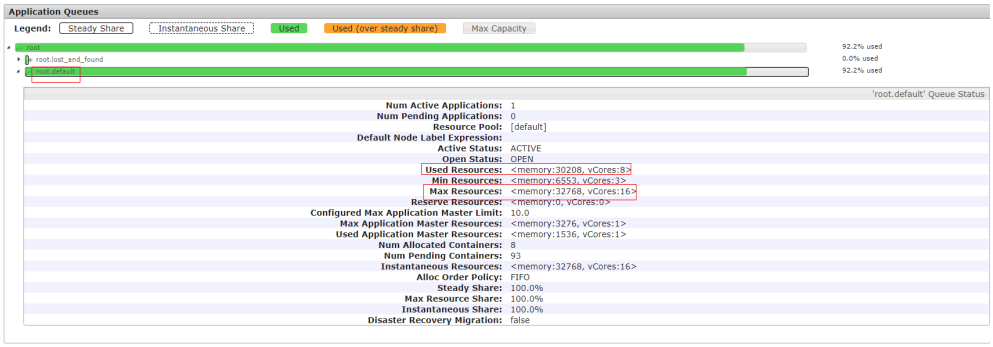
Step 4 Evaluate the expected task execution time based on the service and compare it with the task timeout interval. If the timeout interval is too short, set the timeout interval (**mapreduce.application.timeout.alarm** or **spark.application.timeout.alarm**) of the client to the task expected execution time. Run the task again and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the queue resources are sufficient.

Step 5 Find the task on the native page and view the queue name of the task. Click **Scheduler** on the left of the native page. On the **Applications Queues** page, find the corresponding queue name and expand the queue details, as shown in the following figure.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

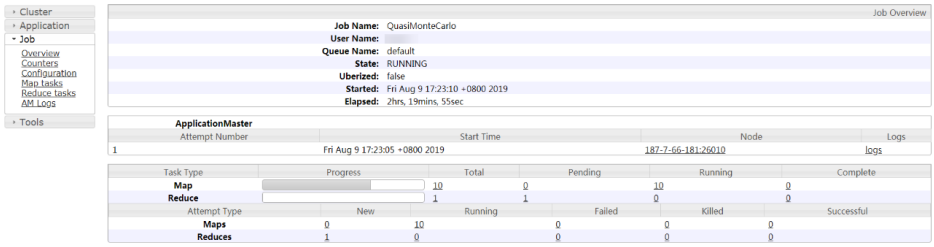


Step 6 Check whether the value of **Used Resources** in the queue details is approximately equal to the value of **Max Resources**, which indicates that the resources in the queue submitted by the task have been used up. If the queue resources are insufficient, choose **Tenant Resources** > **Dynamic Resource Plan** > **Resource Distribution Policy** on FusionInsight Manager and increase the value of **Max Resources** for the queue. Run the task again and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check whether data skew occurs.

Step 7 On the native Yarn page, click *task ID* (for example, **application_1565337919723_0002**) > **Tracking URL:ApplicationMaster** > **job_1565337919723_0002**. The following page is displayed.



Step 8 Choose **Job** > **Map tasks** or **Job** > **Reduce tasks** on the left and check whether the execution time of each Map or Reduce task differs greatly. If yes, task data skew occurs. In this case, you need to balance the task data.

Step 9 Rectify the fault based on the preceding causes and perform the tasks again. Then, check whether the alarm persists.

- If yes, go to **Step 10**.
- If no, no further action is required.

Collect the fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **Yarn** for the target cluster.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.173 ALM-18021 MapReduce Service Unavailable

Description

The alarm module checks the MapReduce service status every 60 seconds. This alarm is generated when the system detects that the MapReduce service is unavailable.

The alarm is cleared when the MapReduce service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18021	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The cluster cannot provide the MapReduce service. For example, MapReduce cannot be used to view task logs or the log archive function is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The JobHistoryServer instance is abnormal.
- The KrbServer service is abnormal.
- The ZooKeeper service abnormal.
- The HDFS service abnormal.
- The Yarn service is abnormal.

Procedure

Check MapReduce service JobHistoryServer instance status.

Step 1 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **MapReduce** > **Instance**.

Step 2 Check whether the running status of JobHistoryServer is **Normal**.

- If yes, go to **Step 11**.
- If no, go to **Step 3**.

Check the KrbServer service status.

Step 3 In the alarm list on FusionInsight Manager, check whether **ALM-25500 KrbServer Service Unavailable** exists.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Rectify the fault by following the steps provided in **ALM-25500 KrbServer Service Unavailable**, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the ZooKeeper service.

Step 5 In the alarm list on FusionInsight Manager, check whether **ALM-13000 ZooKeeper Service Unavailable** exists.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 Rectify the fault by following the steps provided in **ALM-13000 ZooKeeper Service Unavailable**, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check the HDFS service status.

Step 7 In the alarm list on FusionInsight Manager, check whether **ALM-14000 HDFS Service Unavailable** exists.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Rectify the fault by following the steps provided in **ALM-14000 HDFS Service Unavailable**, and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check the Yarn service status.

Step 9 In the alarm list on FusionInsight Manager, check whether **ALM-18000 Yarn Service Unavailable** exists.

- If yes, go to **Step 10**
- If no, go to **Step 11**.

Step 10 Rectify the fault by following the steps provided in **ALM-18000 Yarn Service Unavailable**, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On the FusionInsight Manager home page of the active cluster, choose **O&M** > **Log** > **Download.**

Step 12 Select **MapReduce** in the required cluster from the **Service.**

Step 13 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.174 ALM-18022 Insufficient Yarn Queue Resources

Description

- The alarm module checks Yarn queue resources every 60 seconds. This alarm is generated when available resources or ApplicationMaster (AM) resources of a queue are insufficient.

  This alarm is cleared when available resources are sufficient.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18022	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter Name	Description
Source	Specifies the cluster for which the alarm is generated.
QueueName	Specifies the queue for which the alarm is generated.
QueueMetric	Specifies the metric of the queue for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

- An application being executed takes longer time.
- An application fails to be executed for a long time after being submitted.

Possible Causes

- NodeManager node resources are insufficient.
- The configured maximum resource capacity of the queue is excessively small.
- The configured maximum AM resource percentage is excessively small.

Procedure

Check NodeManager resources.

Step 1 On the FusionInsight Manager, choose **O&M** > **Alarm > Alarms**.

Step 2 View location information of this alarm and check whether **QueueName** is **root** and **QueueMetric** is **Memory** or **QueueName** is **root** and **QueueMetric** is **vCores**.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 The memory or CPU of the Yarn cluster is insufficient. In this case, log in to the node where NodeManager resides and run the **free -g** and **cat /proc/cpuinfo** commands to query the available memory and available CPU of the node, respectively. On FusionInsight Manager, increase the values of **yarn.nodemanager.resource.memory-mb** and **yarn.nodemanager.resource.cpu-vcores** for the Yarn NodeManager based on the query results. Then, restart the NodeManager instance. Check whether the alarm is cleared.

NOTE

During NodeManager restart, containers submitted to this node may be retried to other nodes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 4**.

Checking the maximum resource capacity of a queue.

Step 4 View location information of this alarm and check whether **QueueName** is **<*Tenant Queue*>** and **QueueMetric** is **Memory**, or **QueueName** is **<*Tenant Queue*>** and **QueueMetric** is **vCores** in **Location**, check whether **available Memory =** or **available vCores =** are included in **Additional Information**.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 The memory or CPU of the tenant queue is insufficient. In this case, choose **Tenant Resources** > **Dynamic Resource Plan > Resource Distribution Policy** and increase the value of **Maximum Capacity**. Then, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Step 6 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **Configurations** > **All Configurations**. Enter the keyword "threshold" and click **ResourceManager**. Adjust the threshold values of the following parameters:

If **Additional Information** contains **available Memory =**, change the value of **yarn.queue.memory.alarm.threshold** to a value smaller than that of **available Memory =** in **Additional Information**.

If **Additional Information** contains **available vCores =**, change the value of **yarn.queue.vcore.alarm.threshold** to a value smaller than that of **available vCores =** in **Additional Information**.

Wait for five minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Checking the maximum AM resource percentage.

Step 7 If **available AmMemory =** or **available AmvCores =** is included in **Additional Information**, ApplicationMaster memory or CPU of the tenant queue is insufficient. In this case, choose **Tenant Resources** > **Dynamic Resource Plan** > **Queue Configuration** and increase the value of **Maximum Am Resource Percent**. Then, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 8 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **Configurations** > **All Configurations**. Enter the keyword "threshold" and click **ResourceManager**. Adjust the threshold values of the following parameters:

If **Additional Information** contains **available AmMemory =**, change the value of **yarn.am.memory.alarm.threshold** to a value smaller than that of **available AmMemory =** in **Additional Information**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If **Additional Information** contains **available AmvCores =**, change the value of **yarn.am.vcore.alarm.threshold** to a value smaller than that of **available AmvCores =** in **Additional Information**.

Wait for five minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 Log in to FusionInsight Manager of the active cluster, and choose **O&M** > **Log** > **Download**.

Step 10 Select **Yarn** in the required cluster from the **Service**.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Reference

None

9.11.175 ALM-18023 Number of Pending Yarn Tasks Exceeds the Threshold

Alarm Description

The alarm module checks the number of pending applications in the Yarn root queue every 60 seconds. The alarm is generated when the number exceeds 60.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18023	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
QueueName	Specifies the queue for which the alarm was generated.
QueueMetric	Specifies the queue metric for which the alarm was generated.

Impact on the System

- It takes long time to end an application.
- A new application cannot run after submission.

Possible Causes

- NodeManager node resources are insufficient.
- The maximum resource capacity of the queue and the maximum AM resource percentage are too small.
- The monitoring threshold is too small.

Handling Procedure

Check NodeManager resources.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **ResourceManager (Active)** to access the ResourceManager web UI.

Step 2 Click **Scheduler** and check whether the root queue resources are used up in **Application Queues**.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Expand the capacity of the NodeManager instance of the Yarn service. After the capacity expansion, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check the maximum queue resource capacity and the maximum AM resource percentage.

Step 4 Check whether the resources of the queue corresponding to the pending task are used up.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 On FusionInsight Manager, choose **Tenant Resources** > **Dynamic Resource Plan** and add resources as required. Check whether the alarms are cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Adjust the monitoring thresholds.

Step 6 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Yarn** > **Applications** > **Pending Applications**, and increase the thresholds as required.

Step 7 Check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **Yarn** for the target cluster.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.176 ALM-18024 Pending Yarn Memory Usage Exceeds the Threshold

Alarm Description

The alarm module checks the pending memory of Yarn every 60 seconds. The alarm is generated when the pending memory exceeds the threshold. Pending memory indicates the total memory that is not allocated to submitted Yarn applications.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18024	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
QueueName	Specifies the queue for which the alarm was generated.
QueueMetric	Specifies the queue metric for which the alarm was generated.

Impact on the System

- It takes long time to end an application.
- A new application cannot run after submission.

Possible Causes

- NodeManager node resources are insufficient.
- The maximum resource capacity of the queue and the maximum AM resource percentage are too small.
- The monitoring threshold is too small.

Handling Procedure

Check NodeManager resources.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **ResourceManager (Active)** to access the ResourceManager web UI.

Step 2 Click **Scheduler** and check whether the root queue resources are used up in **Application Queues**.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Expand the capacity of the NodeManager instance of the Yarn service. After the capacity expansion, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check the maximum queue resource capacity and the maximum AM resource percentage.

Step 4 Check whether the resources of the queue corresponding to the pending task are used up.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 On FusionInsight Manager, choose **Tenant Resources** > **Dynamic Resource Plan** and add resources as required. Check whether the alarms are cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Adjust the monitoring thresholds.

Step 6 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Yarn** > **CPU and Memory** > **Pending Memory**, and increase the threshold as required.

Step 7 Check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **Yarn** for the target cluster.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.177 ALM-18025 Number of Terminated Yarn Tasks Exceeds the Threshold

Description

The alarm module checks the number of terminated applications in the Yarn root queue every 60 seconds. The alarm is generated when the number exceeds 50 for three consecutive times.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18025	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Cluster Name	Specifies the cluster for which the alarm is generated.
Service Name	Specifies the service for which the alarm is generated.
Role Name	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A large number of application tasks are forcibly terminated.

Possible Causes

- The user forcibly terminates a large number of tasks.
- The system terminates tasks due to some error.

Procedure

Check the alarm details.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** to go to the alarm page.

Step 2 View **Additional Information** in the alarm details to check whether the alarm threshold is too small.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Yarn** > **Other** > **Terminated Applications of root queue** to modify the threshold. Go to **Step 6**.

Step 4 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **ResourceManager(Active)** to access the ResourceManager web UI.

Step 5 Click **KILLED** in **Applications** and click the task on the top. View the description of **Diagnostics** and rectify the fault based on the task termination details (for example, the task is terminated by a user).

Step 6 Wait for 3 minutes and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On the FusionInsight Manager, choose **O&M > Log > Download**.

Step 8 Expand the **Service** drop-down list, and select **Yarn** for the target cluster.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.178 ALM-18026 Number of Failed Yarn Tasks Exceeds the Threshold

Description

The alarm module checks the number of failed applications in the Yarn root queue every 60 seconds. The alarm is generated when the number exceeds 50 for three consecutive times.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18026	Major	Yes

Parameters

Name	Meaning
Cluster Name	Specifies the cluster for which the alarm is generated.
Service Name	Specifies the service for which the alarm is generated.
Role Name	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

- A large number of application tasks fail to be executed.
- Failed tasks need to be submitted again.

Possible Causes

The task fails to be executed due to some error.

Procedure

Check the alarm details.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** to go to the alarm page.

Step 2 View **Additional Information** in the alarm details to check whether the alarm threshold is too small.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Yarn** > **Other** > **Failed Applications of root queue** to modify the threshold. Go to **Step 6**.

Step 4 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **ResourceManager(Active)** to access the ResourceManager web UI.

Step 5 Click **FAILED** in **Applications** and click the task on the top. View the description of **Diagnostics** and rectify the fault based on the task failure causes.

Step 6 Wait for 3 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On the FusionInsight Manager, choose O&M > Log > Download.

Step 8 Expand the **Service** drop-down list, and select **Yarn** for the target cluster.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.179 ALM-18027 JobHistoryServer Process Is Abnormal

Alarm Description

The JobHistoryServer process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18027	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Run the command in **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check the process state:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.mapreduce.v2.hs.JobHistoryServer | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **Mapreduce** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.180 ALM-18028 TimeLineServer Process Is Abnormal

Alarm Description

The TimeLineServer process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
18028	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Run the command in **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check the process state:

ps ww -eo stat,cmd| grep -w org.apache.hadoop.yarn.server.applicationhistoryservice.ApplicationHistoryServer | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **Yarn** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.181 ALM-19000 HBase Service Unavailable

Alarm Description

The alarm module checks the HBase service status every 120 seconds. This alarm is generated when the HBase service is unavailable.

This alarm is cleared when the HBase service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19000	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Operations such as data read/write and table creation cannot be performed.

Possible Causes

- ZooKeeper is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- HDFS is abnormal.
- HBase is abnormal.
- The network connection is abnormal.
- The service configuration value is incorrect.

Handling Procedure

Check the ZooKeeper service status.

Step 1 In the service list on FusionInsight Manager, check whether **Running Status** of ZooKeeper is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 In the alarm list, check whether **ALM-13000 ZooKeeper Service Unavailable** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by performing the operations provided for **ALM-13000 ZooKeeper Service Unavailable**.

Step 4 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the HDFS service status.

Step 5 In the alarm list, check whether **ALM-14000 HDFS Service Unavailable** exists.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Rectify the fault by performing the operations provided for **ALM-14000 HDFS Service Unavailable**.

Step 7 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 8 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **HDFS**. On the **Dashboard**, check whether the **Safe Mode** is **ON** for HDFS.

- If yes, go to **Step 9**.
- If no, go to **Step 12**.

Step 9 Log in to the node where the HDFS client is deployed as the **root** user.

Run the following command to go to the client installation directory and load environment variables:

```
cd Client installation directory
source bigdata_env
```

If Kerberos authentication is enabled for the cluster, complete security authentication:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
kinit Username of the service user who has the HDFS operation permissions
```

Enter the required password.

Step 10 Run the following command to manually exit the safe mode:

```
hdfs dfsadmin -safemode leave
```

Step 11 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check the HBase service status.

Step 12 On FusionInsight Manager, choose **Cluster** > **Services** > **HBase**.

Step 13 Check whether there is one active HMaster and one standby HMaster.

- If yes, go to **Step 15**.
- If no, go to **Step 14**.

Step 14 Click **Instances** and select the HMaster instance whose status is not **Active**. Click **More** and select **Restart Instance** to restart HMaster. Then check whether there is one active HMaster and one standby HMaster.

- If yes, go to **Step 15**.
- If no, go to **Step 21**.

NOTICE

During the HMaster restart, table operations cannot be performed, and the HBase web UI is inaccessible. Data read and write operations are not affected.

Step 15 Choose **Cluster** > **Services** > **HBase** and click **HMaster(Active)** to access the HMaster WebUI.

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 16 Check whether at least one RegionServer exists under **Region Servers**.

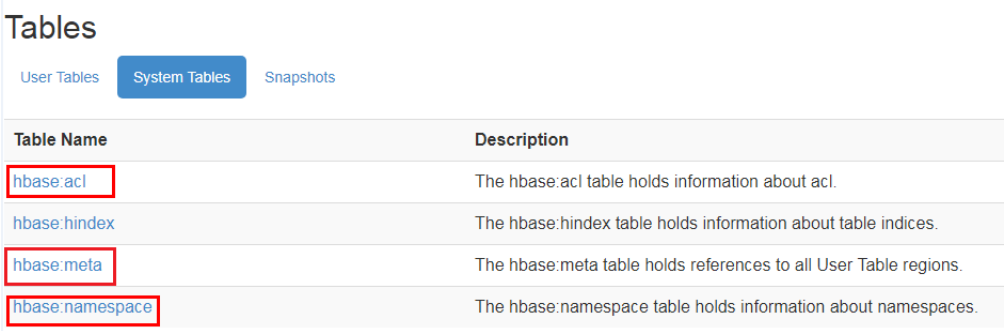
- If yes, go to **Step 17**.
- If no, go to **Step 21**.

Step 17 Choose **Tables** > **System Tables** and check whether **hbase:meta**, **hbase:namespace**, and **hbase:acl** exist in the **Table Name** column, as shown in **Figure 9-8**.

- If yes, go to **Step 18**.
- If no, go to **Step 19**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-8 HBase system tables



Step 18 Click **hbase:meta**, **hbase:namespace**, and **hbase:acl** to check whether all pages can be opened. If all of them can be opened, the tables are normal.

- If yes, go to **Step 19**.
- If no, go to **Step 25**.

NOTE

In a normal cluster, ACL permission control is disabled for HBase by default. The **hbase:acl** table is generated only after ACL permission control is manually enabled. In this case, you need to check this table.

Step 19 View the HMaster startup status.

On the **Tasks** page shown in **Figure 9-9**, the **RUNNING** value in the **State** column indicates that HMaster is being started and provides how much time HMaster keeps in that state. As shown in **Figure 9-10**, if the state is **COMPLETE**, HMaster has been started.

Check whether HMaster has been in the **RUNNING** state for a long time.

Figure 9-9 HMaster being started



Figure 9-10 HMaster startup completed

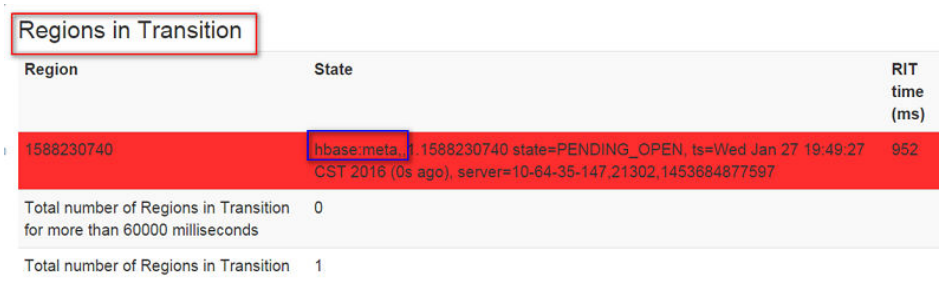


- If yes, go to **Step 20**.
- If no, go to **Step 21**.

Step 20 On the HMaster web UI, check whether any **hbase:meta** is in the **Regions in Transition** state for a long time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-11 Regions in Transition



- If yes, go to **Step 21**.
- If no, go to **Step 22**.

Step 21 After ensuring that services are not affected, log in to FusionInsight Manager, choose **Cluster**. Click the name of the required cluster, choose **Services** > **HBase**, click **More**, and select **Restart Service**. In the dialog box that is displayed, enter the password, and click **OK**.

- If yes, go to **Step 22**.
- If no, go to **Step 25**.

NOTICE

During HBase service restart, the service is unavailable. For example, data cannot be read or written, table operations cannot be performed, and the HBase web UI is inaccessible.

Step 22 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 25**.

Check whether the HBase configurations are correctly modified.

Step 23 On FusionInsight Manager, choose **Audit**. On the **Audit** page, click **Advanced Search**, click ··· on the right of **Operation Type**, select **Save configuration**, click **OK**, and click **Search**.

Step 24 In the search result, check whether the historical configurations of HBase-related services in the **Service** column, such as ZooKeeper, HDFS, and HBase, may affect the HBase service status. **Table 9-122** lists some configurations that may affect the HBase service status.

Table 9-122 Configurations affecting the HBase service status

Parameter	Possible Impact
GC_OPTS	The memory configuration may be improper. You need to check the health status of instance processes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Possible Impact
hbase.rpc.protection	If the HBase service is not restarted offline after the value of this parameter is changed, the connection authentication fails and the HBase service becomes abnormal.
hbase.regionserver.metahandler.count	If there are too many regions in the cluster but this parameter is set to a small value, RIT may occur and regions cannot be brought online for a long time.
hbase.regionserver.thread.compaction.large	If this parameter is set to a large value, the node CPU usage may be too high.
hbase.regionserver.thread.compaction.small	If this parameter is set to a large value, the node CPU usage may be too high.
hbase.coprocessor.master.classes	If a custom coprocessor is used in the configuration, a logic error may cause the service to be unavailable.
hbase.coprocessor.region.classes	If a custom coprocessor is used in the configuration, a logic error may cause the service to be unavailable.
hbase.coprocessor.regionserver.classes	If a custom coprocessor is used in the configuration, a logic error may cause the service to be unavailable.
zookeeper.session.timeout	If this parameter is set to a small value, the connection between HBase and ZooKeeper times out too quickly. As a result, the HMaster instance and RegionServer may restart repeatedly.

Check the network connection between HMaster and dependent components.

Step 25 On FusionInsight Manager, choose **Cluster** > **Services** > **HBase**.

Step 26 Click **Instances**. In the HMaster instance list, record the management IP address of the active HMaster instance.

Step 27 Log in to the active HMaster node as user **omm** using the IP address obtained in **Step 26**.

Step 28 Run the **ping** command to check whether the network connection between the active HMaster node and the host where the dependent components reside is normal. (The dependent components include ZooKeeper, HDFS, and Yarn. The method of obtaining the IP address of the host where the dependent components reside is the same as that of obtaining the IP address of the active HMaster node.)

- If yes, go to **Step 31**.
- If no, go to **Step 29**.

Step 29 Contact the network administrator to restore the network.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 30 In the alarm list, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 31**.

Collect fault information.

Step 31 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 32 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select the following services for the target cluster:

- ZooKeeper
- HDFS
- HBase

Step 33 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 34 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.182 ALM-19006 HBase Replication Sync Failed

Description

The alarm module checks the HBase DR data synchronization status every 30 seconds. When disaster recovery (DR) data fails to be synchronized to a standby cluster, the alarm is triggered.

When DR data synchronization succeeds, the alarm is cleared.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
19006	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

HBase data in the cluster cannot be synchronized to the standby cluster. Synchronization data is stacked, causing a large amount of active/standby data inconsistency. As a result, the latest data cannot be read from the standby cluster after an active/standby DR switchover or dual-read. If the alarm persists, the storage space of the primary cluster and ZooKeeper nodes will be stacked, leading to service faults in the primary cluster.

Possible Causes

- The HBase service on the standby cluster is abnormal.
- A network exception occurs.

Procedure

Observe whether the system automatically clears the alarm.

Step 1 On the FusionInsight Manager portal of the active cluster, click **O&M** > **Alarm** > **Alarms.**

Step 2 In the alarm list, click the alarm to obtain alarm generation time from **Generated** of the alarm. Check whether the alarm has existed for five minutes.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 Wait five minutes and check whether the system automatically clears the alarm.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the HBase service status of the standby cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Log in to the FusionInsight Manager portal of the active cluster, and click **O&M** > **Alarm** > **Alarms.**

Step 5 In the alarm list, click the alarm to obtain **HostName** from **Location**.

Step 6 Access the node where the HBase client of the active cluster resides as user **omm**.

If the cluster uses a security mode, perform security authentication first and then access the **hbase shell** interface as user **hbase**.

cd /opt/client

source ./bigdata_env

kinit *hbaseuser*

Step 7 Run the **status 'replication', 'source'** command to check the DR synchronization status of the faulty node.

The DR synchronization status of a node is as follows.

```
10-10-10-153:
SOURCE: PeerID=abc, SizeOfLogQueue=0, ShippedBatches=2, ShippedOps=2, ShippedBytes=320,
LogReadInBytes=1636, LogEditsRead=5, LogEditsFiltered=3, SizeOfLogToReplicate=0,
TimeForLogToReplicate=0, ShippedHFiles=0, SizeOfHFileRefsQueue=0, AgeOfLastShippedOp=0,
TimeStampsOfLastShippedOp=Mon Jul 18 09:53:28 CST 2016, Replication Lag=0,
FailedReplicationAttempts=0
SOURCE: PeerID=abc1, SizeOfLogQueue=0, ShippedBatches=1, ShippedOps=1, ShippedBytes=160,
LogReadInBytes=1636, LogEditsRead=5, LogEditsFiltered=3, SizeOfLogToReplicate=0,
TimeForLogToReplicate=0, ShippedHFiles=0, SizeOfHFileRefsQueue=0, AgeOfLastShippedOp=16788,
TimeStampsOfLastShippedOp=Sat Jul 16 13:19:00 CST 2016, Replication Lag=16788,
FailedReplicationAttempts=5
```

Step 8 Obtain **PeerID** corresponding to a record whose **FailedReplicationAttempts** value is greater than 0.

In the preceding step, data on the faulty node 10-10-10-153 fails to be synchronized to a standby cluster whose **PeerID** is **abc1**.

Step 9 Run the **list_peers** command to find the cluster and the HBase instance corresponding to the **PeerID** value.

```
PEER_ID CLUSTER_KEY STATE TABLE_CFS
abc1 10.10.10.110,10.10.10.119,10.10.10.133:2181:/hbase2 ENABLED
abc 10.10.10.110,10.10.10.119,10.10.10.133:2181:/hbase ENABLED
```

In the preceding information, **/hbase2** indicates that data is synchronized to the HBase2 instance of the standby cluster.

Step 10 In the service list of FusionInsight Manager of the standby cluster, check whether the running status of the HBase instance obtained by using **Step 9** is **Normal**.

- If yes, go to **Step 14**.
- If no, go to **Step 11**.

Step 11 In the alarm list, check whether the **ALM-19000 HBase Service Unavailable** alarm is generated.

- If yes, go to **Step 12**.
- If no, go to **Step 14**.

Step 12 Follow troubleshooting procedures in **ALM-19000 HBase Service Unavailable** to rectify the fault.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Wait for a few minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Check network connections between RegionServers on active and standby clusters.

Step 14 Log in to the FusionInsight Manager portal of the active cluster, and click **O&M** > **Alarm** > **Alarms.**

Step 15 In the alarm list, click the alarm to obtain **HostName** from **Location**.

Step 16 Use the IP address obtained in **Step 15** to log in to a faulty RegionServer node as user **omm**.

Step 17 Run the **ping** command to check whether network connections between the faulty RegionServer node and the host where RegionServer of the standby cluster resides are in the normal state.

- If yes, go to **Step 20**.
- If no, go to **Step 18**.

Step 18 Contact the network administrator to restore the network.

Step 19 After the network is running properly, check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 20**.

Collect fault information.

Step 20 On the FusionInsight Manager interface of active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 21 In the **Service** drop-down list box, select **HBase** in the required cluster.

Step 22 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 23 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.183 ALM-19007 HBase GC Time Exceeds the Threshold

Alarm Description

The system checks the old generation garbage collection (GC) time of the HBase service every 60 seconds. This alarm is generated when the detected old generation GC time exceeds the threshold (exceeds 5 seconds for three consecutive checks by default). To change the threshold, on the FusionInsight Manager homepage, choose **O&M** > **Alarm** > **Thresholds** > **HBase** > **GC** > **GC time for old generation**. This alarm is cleared when the old generation GC time of the HBase service is shorter than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19007	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.

Impact on the System

If the GC time of the old generation exceeds the threshold, the read and write of HBase data will slow down. In severe cases, the request times out.

Possible Causes

The memory of HBase instances is overused, the heap memory is inappropriately allocated, or a large number of I/O operations exist in HBase. As a result, GCs occur frequently.

Handling Procedure

Check the GC time.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On the FusionInsight Manager homepage, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose **ID** is **19007**. Then check the role name in **Location** and confirm the IP address of the instance.

- If the role for which the alarm is generated is HMaster, go to **Step 2**.
- If the role for which the alarm is generated is RegionServer, go to **Step 3**.

Step 2 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HBase** > **Instances** and click the HMaster for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **GC** > **Garbage Collection (GC) Time of HMaster** and click **OK** to check whether the value of **GC time for old generation** is greater than the threshold (exceeds 5 seconds for three consecutive checks periods by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 3 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HBase** > **Instances** and click the RegionServer for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **GC** > **Garbage Collection (GC) Time of RegionServer** and click **OK** to check whether the value of **GC time for old generation** is greater than the threshold (exceeds 5 seconds for three consecutive checks periods by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Check the current JVM configuration.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **HBase** > **Configurations** > **All Configurations**. In Search, enter **GC_OPTS** to check the **GC_OPTS** memory parameter of role HMaster(HBase->HMaster), RegionServer(HBase->RegionServer). Adjust the values of **-Xmx** and **-XX:CMSInitiatingOccupancyFraction** of the GC_OPTS parameter by referring to the Note.

1. Suggestions on GC parameter configurations for HMaster
   - Set **-Xms** and **-Xmx** to the same value to prevent JVM from dynamically adjusting the heap memory size and affecting performance.
   - Set **-XX:NewSize** to the value of **-XX:MaxNewSize**, which is one eighth of **-Xmx**.
   - For large-scale HBase clusters with a large number of regions, increase values of **GC_OPTS** parameters for HMaster. Specifically, set **-Xmx** to 4 GB if the number of regions is less than 100,000. If the number of regions is more than 100,000, set **-Xmx** to be greater than or equal to 6 GB. For each increased 35,000 regions, increase the value of **-Xmx** by 2 GB. The maximum value of **-Xmx** is 32 GB.
2. Suggestions on GC parameter configurations for RegionServer
   - Set **-Xms** and **-Xmx** to the same value to prevent JVM from dynamically adjusting the heap memory size and affecting performance.
   - Set **-XX:NewSize** to one eighth of **-Xmx**.
   - Set the memory for RegionServer to be greater than that for HMaster. If sufficient memory is available, increase the heap memory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - Set **-Xmx** based on the machine memory size. Specifically, set **-Xmx** to 32 GB if the machine memory is greater than 200 GB, to 16 GB if the machine memory is greater than 128 GB and less than 200 GB, and to 8 GB if the machine memory is less than 128 GB. When **-Xmx** is set to 32 GB, a RegionServer node supports 2000 regions and 200 hotspot regions.
  - **XX:CMSInitiatingOccupancyFraction** to be less than and equal to **85**, and it is calculated as follows: 100 x (hfile.block.cache.size + hbase.regionserver.global.memstore.size)

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager interface of active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 7 In the **Service** drop-down list box, select **HBase** in the required cluster.

Step 8 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 9 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.184 ALM-19008 Heap Memory Usage of the HBase Process Exceeds the Threshold

Description

The system checks the HBase service status every 30 seconds. The alarm is generated when the heap memory usage of an HBase service exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
19008	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The available HBase memory is insufficient, which may cause node restart. During the node restart, the read/write request delay on the node increases or fails.

Possible Causes

The heap memory of the HBase service is overused or the heap memory is inappropriately allocated.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, click **O&M** > **Alarm** > **Alarms** and select the alarm whose **ID** is **19008**. Then check the role name in **Location** and confirm the IP address of the instance.

- If the role for which the alarm is generated is HMaster, go to **Step 2**.
- If the role for which the alarm is generated is RegionServer, go to **Step 3**.

Step 2 On the FusionInsight Manager, choose **Cluster** > **Services** > **HBase** > **Instances** and click the HMaster for which the alarm is generated. On the displayed **Dashboard** page, click the drop-down arrow in the upper right corner of the **Chart** area, choose **Customize** > **CPU and Memory** > **HMaster Heap Memory Usage and Direct Memory Usage Statistics**, and click **OK**.

Check whether the HBase service process has consumed 90% of its allocated maximum heap memory.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 3 On the FusionInsight Manager, choose **Cluster** > **Services** > **HBase** > **Instances** and click the RegionServer for which the alarm is generated. On the displayed **Dashboard** page, click the drop-down arrow in the upper right corner of the **Chart** area, choose **Customize** > **CPU and Memory** > **RegionServer Heap Memory Usage and Direct Memory Usage Statistics**, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the HBase service process has consumed 90% of its allocated maximum heap memory.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager, choose **Cluster** > **Services** > **HBase** > **Configurations**, and click **All Configurations**. Choose **HMaster/RegionServer** > **System**. Increase the value of **-Xmx** in **GC_OPTS** by referring to the Note.

NOTE

- Recommended GC parameter settings for HMaster
  - Set **-Xms** and **-Xmx** to the same value to prevent JVM from dynamically adjusting the heap memory size and affecting performance.
  - Set **-XX:NewSize** to the value of **-XX:MaxNewSize**, which is one eighth of **-Xmx**.
  - For large-scale HBase clusters with a large number of regions, increase values of **GC_OPTS** parameters for HMaster. Specifically, set **-Xmx** to 4 GB if the number of regions is less than 100,000. If the number of regions is more than 100,000, set -Xmx to be greater than or equal to 6 GB. For each increased 35,000 regions, increase the value of **-Xmx** by 2 GB. The maximum value of **-Xmx** is 32 GB.
- Recommended GC parameter settings for RegionServer
  - Set **-Xms** and **-Xmx** to the same value to prevent JVM from dynamically adjusting the heap memory size and affecting performance.
  - Set **-XX:NewSize** to one eighth of **-Xmx**.
  - Set the memory for RegionServer to be greater than that for HMaster. If sufficient memory is available, increase the heap memory.
  - Set **-Xmx** based on the machine memory size. Specifically, set **-Xmx** to 32 GB if the machine memory is greater than 200 GB, to 16 GB if the machine memory is greater than 128 GB and less than 200 GB, and to 8 GB if the machine memory is less than 128 GB. When **-Xmx** is set to 32 GB, a RegionServer node supports 2000 regions and 200 hotspot regions.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **HBase** in the required cluster from the **Service** drop-down list.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.185 ALM-19009 Direct Memory Usage of the HBase Process Exceeds the Threshold

Description

The system checks the HBase service status every 30 seconds. The alarm is generated when the direct memory usage of an HBase service exceeds the threshold (90% of the maximum memory).

The alarm is cleared when the direct memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
19009	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.

Impact on the System

The available HBase direct memory is insufficient, which may cause node restart. During the node restart, the read/write request delay on the node increases or fails.

Possible Causes

The direct memory of the HBase service is overused or the direct memory is inappropriately allocated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check direct memory usage.

Step 1 On the FusionInsight Manager portal, click **O&M** > **Alarm** > **Alarms** and select the alarm whose **ID** is **19009**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

- If the role for which the alarm is generated is HMaster, go to **Step 2**.
- If the role for which the alarm is generated is RegionServer, go to **Step 3**.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase** > **Instance** and click the HMaster for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory** > **HMaster Heap Memory Usage and Direct Memory Usage Statistics** and click **OK** to check whether the used direct memory of the HBase service reaches 90% of the maximum direct memory specified for HBase.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 3 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase** > **Instance** and click the RegionServer for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **CPU and Memory** > **RegionServer Heap Memory Usage and Direct Memory Usage Statistics** and click **OK** to check whether the used direct memory of the HBase service reaches 90% of the maximum direct memory specified for HBase.

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase** > **Configurations**, and click **All Configurations**. Choose **HMaster/RegionServer** > **System** and check whether **XX:MaxDirectMemorySize** exists in **GC_OPTS**.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase** > **Configurations**, and click **All Configurations**. Choose **HMaster/RegionServer** > **System** and delete **XX:MaxDirectMemorySize** from **GC_OPTS**.

Step 6 Check whether the **ALM-19008 Heap Memory Usage of the HBase Process Exceeds the Threshold** alarm is generated.

- If yes, handle the alarm by referring to **ALM-19008 Heap Memory Usage of the HBase Process Exceeds the Threshold**.
- If no, go to **Step 8**.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 On the FusionInsight Manager interface of active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 9 In the **Service** in the required cluster drop-down list box, select **HBase**.

Step 10 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.186 ALM-19011 RegionServer Region Number Exceeds the Threshold

Alarm Description

The system checks the number of regions on each RegionServer in each HBase service instance every 30 seconds. The region number is displayed on the HBase service monitoring page and RegionServer role monitoring page. This alarm is generated when the number of regions on a RegionServer exceeds the threshold (default value: 2000) for 20 consecutive times. The threshold can be changed by choosing **O&M** > **Alarm** > **Thresholds** > **HBase**. This alarm is cleared when the number of regions is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19011	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the number of RegionServer regions exceeds the threshold, too many Regions increase the load of RegionServer, causing resource bottlenecks such as memory, disk I/O, and CPU. As a result, request response becomes slow or even times out.

Possible Causes

- The RegionServer region distribution is unbalanced.
- The HBase cluster scale is too small.

Handling Procedure

View alarm location information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **19011**, and view the service instance and host name in **Location**.

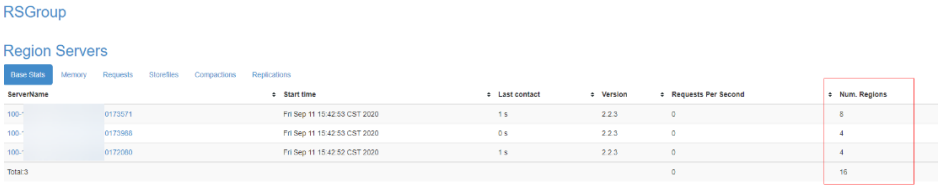
Step 2 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HBase**. Click **HMaster(Active)** to go to the WebUI of the HBase instance and check whether regions are evenly distributed on the RegionServer.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

- If yes, go to **Step 9**.
- If no, go to **Step 3**.

Figure 9-12 WebUI of HBase instance



Enable load balancing.

Step 3 Log in to the node where the HBase client is deployed as the **root** user.

Run the following command to open the client installation directory:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
cd Client installation directory
```

Configure environment variables:

```
source bigdata_env
```

If the cluster uses the security mode, perform security authentication.

```
kinit Service user who has the HBase operation permissions
```

Enter the required password.

Step 4 Run the following command to go to the HBase shell CLI:

```
hbase shell
```

Check whether the load balancing function is enabled:

```
balancer_enabled
```

- If yes, go to **Step 6**.
- If no, go to **Step 5**.

Step 5 In the HBase shell CLI, run the following command to enable the load balancing function:

```
balance_switch true
```

Verify that the function is enabled:

```
balancer_enabled
```

Step 6 Run the following command to manually trigger load balancing:

```
balancer
```

Enable and manually trigger load balancing during off-peak hours.

Step 7 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HBase**, and click **HMaster(Active)**. On the displayed WebUI of the HBase instance, refresh the page and check whether regions are evenly distributed.

- If yes, go to **Step 8**.
- If no, go to **Step 21**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Delete unnecessary HBase tables.

Step 9 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HBase** and click **HMaster(Active)**. On the displayed WebUI of the HBase instance, view tables stored in the HBase service instance and record unnecessary tables that can be deleted.

Step 10 In the HBase shell CLI, run **disable** and **drop** to delete the unnecessary tables to reduce regions.

```
disable 'Name of the table to be deleted'
drop 'Name of the table to be deleted'
```

WARNING

Exercise caution to ensure data is deleted correctly.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 In the HBase shell CLI, run the following command to check whether the load balancing function is enabled:

```
balancer_enabled
```

- If yes, go to **Step 13**.
- If no, go to **Step 12**.

Step 12 In the HBase shell CLI, run the following command to enable the load balancing function:

```
balance_switch true
```

Verify that the function is enabled:

```
balancer_enabled
```

Step 13 In the HBase shell CLI, run the following command to manually trigger the load balancing function:

```
balancer
```

Step 14 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **HBase** and click **HMaster(Active)**. Click the HBase service instance that the alarm was generated for. On the displayed WebUI, refresh the page and check whether regions are evenly distributed.

- If yes, go to **Step 15**.
- If no, go to **Step 21**.

Step 15 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Adjust the threshold.

Step 16 On the FusionInsight Manager homepage, choose **O&M** > **Alarm** > **Thresholds** > **HBase** > **Regions(RegionServer)**, select the applied rule, and click **Modify** to check whether the threshold is proper.

- If it is excessively small, increase the threshold as required and go to **Step 17**.
- If it is proper, go to **Step 18**.

Step 17 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Perform system capacity expansion.

Step 18 Add nodes to the HBase cluster and add RegionServer instances to the nodes. Then enable and manually trigger the load balancing function.

Step 19 On the FusionInsight Manager homepage, choose **Cluster** > **Services** and click **HMaster(Active)**. Click the HBase service instance that the alarm was generated for. On the displayed WebUI, refresh the page and check whether regions are evenly distributed.

- If yes, go to **Step 20**.
- If no, go to **Step 21**.

Step 20 Check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 21**.

Collect fault information.

Step 21 On the FusionInsight Manager home page of the active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 22 Select **HBase** in the required cluster from the **Service**.

Step 23 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 24 Contact O&M personnel and send the collected fault logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.187 ALM-19012 HBase System Table Directory or File Lost

Description

The system checks whether HBase directories and files exist on the HDFS every 120 seconds. This alarm is generated when the system detects that the files or directories do not exist. This alarm is cleared when the files or directories are restored.

Check the following items:

- Directory of the namespace **hbase** on the HDFS
- **hbase.version** file
- Directory of the table **hbase:meta** on the HDFS, .tableinfo file, and .regioninfo file
- Directory of the table **hbase:namespace** on the HDFS, .tableinfo file, and .regioninfo file
- Directory of the table **hbase:hindex** on the HDFS, .tableinfo file, and .regioninfo file
- Directory of the **hbase:acl** table on the HDFS, .tableinfo, and .regioninfo file (This table does not exist in the common mode cluster by default.)

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
19012	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The HBase service fails to be restarted or started. As a result, all HBase service requests fail.

Possible Causes

Files or directories on the HDFS are missing.

Procedure

Locate the alarm cause.

Step 1 On the FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Click this alarm whose ID is **19012** and check whether **Alarm Cause** indicates unknown errors.

- If yes, go to **Step 4**.
- If no, go to **Step 2**

Step 2 On the FusionInsight Manager home page, choose **O&M** > **Backup and Restoration** > **Backup Management**. Check whether there are success records of the backup task named **default** or other HBase metadata backup tasks that have been successfully executed.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Use the latest backup metadata to restore the metadata of the HBase service.

Collect fault information.

Step 4 On the FusionInsight Manager page of the active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 5 In the **Service** area, select faulty HBase services in the required cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.188 ALM-19013 Duration of Regions in transaction State Exceeds the Threshold

Description

The system checks the number of regions in transaction state on HBase every 300 seconds. This alarm is generated when the system detects that the duration of regions in transaction state exceeds the threshold for two consecutive times. This alarm is cleared when all timeout regions are restored.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
19013	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Some data in the table gets lost or becomes unavailable.

Possible Causes

- Compaction is permanently blocked.
- The HDFS files are abnormal.

Procedure

Locate the alarm cause.

Step 1 On the FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, select this alarm, and view the **HostName** and **RoleName** in **Location**.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase**, Click the drop-down menu in the chart area and choose **Customize** > **Service** > **Region in transaction count**. Click **OK** and view **Region in transaction count over threshold**. (The default threshold is 60 seconds.)

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase** > **HMaster (Active)** > **Tables** to check whether the regions of only one table transaction status time out.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 Run the **hbase hbck** command on the client and check whether the error message "No table descriptor file under hdfs://hacluster/hbase/data/default/table" is displayed.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Log in to the client as user **root**. Run the following commands:

cd *client installation directory*

source bigdata_env

If the cluster is in security mode, run the **kinit hbase** command

Log in to the HMaster WebUI, choose **Procedure & Locks** in the navigation tree, and check whether any process ID is in the **Waiting** state in **Procedures**. If yes, run the following command to release the procedure lock:

hbase hbck -j *client installation directory*/**HBase/hbase/tools/hbase-hbck2-*.jar bypass -o** *pid*

Check whether the state is in the **Bypass** state. If the procedure on the UI is always in **RUNNABLE(Bypass)** state, perform an active/standby switchover. Run the **assigns** command to bring the region online again.

hbase hbck -j *client installation directory*/**HBase/hbase/tools/hbase-hbck2-*.jar assigns -o** *regionName*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Run the **hbase hbck** command on the client and check whether the error message "No table descriptor file under hdfs://hacluster/hbase/data/default/table" is displayed.

- If yes, go to **Step 7**.
- If no, no further action is required.

Collect fault information.

Step 7 On the FusionInsight Manager page of the active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 8 In the **Service** area, select faulty HBase services in the required cluster.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.189 ALM-19014 Capacity Quota Usage on ZooKeeper Exceeds the Threshold Severely

Alarm Description

The system checks the ZNode usage of the HBase service every 120 seconds. This alarm is generated when the ZNode capacity usage of the HBase service exceeds the critical alarm threshold (90% by default).

This alarm is cleared when the ZNode capacity usage is less than the critical alarm threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19014	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Threshold	Specifies the threshold for generating the alarm.

Impact on the System

This alarm indicates that the capacity usage of the ZNode of HBase has exceeded the threshold severely. As a result, the write request of the HBase service fails.

Possible Causes

- DR is configured for HBase, and data synchronization fails or is slow in DR.
- A large number of WAL files are being split in the HBase cluster.

Handling Procedure

Check the capacity configuration and usage of ZNodes.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, select the alarm whose ID is **19014**, and view the threshold in **Additional Information**.

Step 2 Log in to the HBase client as user **root**. Run the following command to go to the client installation directory:

cd *Client installation directory*

Run the following command to set environment variables:

source bigdata_env

If the cluster uses the security mode, run the following command to perform security authentication:

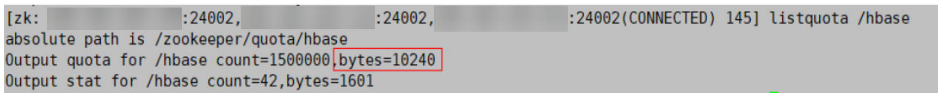
kinit hbase

Enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 3 Run the **hbase zkcli** command to log in to the ZooKeeper client and run the **listquota /hbase** command to check the ZNode capacity quota of the HBase service. The ZNode root directory in the command is specified by the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

zookeeper.znode.parent parameter of the HBase service. The marked area in the following figure shows the capacity configuration of the root ZNode of the HBase service.



Step 4 Run the **getusage /hbase/splitWAL** command to check the capacity usage of the ZNode. Check whether the ratio of **Data size** to the ZNode capacity quota is close to the alarm threshold.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **12007**, **19000**, or **19013** and the **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 6 Run the **getusage /hbase/replication** command to check the capacity usage of the ZNode. Check whether the ratio of **Data size** to the ZNode capacity quota is close to the alarm threshold.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **19006** and **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether the alarm is cleared five minutes later.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect the fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **HBase** for the target cluster.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.190 ALM-19015 Quantity Quota Usage on ZooKeeper Exceeds the Threshold

Alarm Description

The system checks the ZNode usage of the HBase service every 120 seconds. This alarm is generated when the system detects that the ZNode quantity usage of the HBase service exceeds the alarm threshold (75% by default).

This alarm is cleared when the ZNode quantity usage is less than the alarm threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19015	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Threshold	Specifies the threshold for generating the alarm.

Impact on the System

This alarm indicates that the Znode quantity usage in the HBase service has exceeded the threshold. If this alarm is not handled in a timely manner, the problem severity may be escalated to **Critical** and data fails to be written.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- DR is configured for HBase, and data synchronization fails or is slow in DR.
- A large number of WAL files are being split in the HBase cluster.

Handling Procedure

Check the quantity quota and usage of ZNodes.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, select the alarm whose ID is **19015**, and view the threshold in **Additional Information**.

Step 2 Log in to the HBase client as user **root**. Run the following command to go to the client installation directory:

cd *Client installation directory*

Run the following command to set environment variables:

source bigdata_env

If the cluster uses the security mode, run the following command to perform security authentication:

kinit hbase

Enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 3 Run the **hbase zkcli** command to log in to the ZooKeeper client and run the **listquota /hbase** command to check the ZNode quantity quota of the HBase service. The ZNode root directory in the command is specified by the **zookeeper.znode.parent** parameter of the HBase service. The marked area in the following figure shows the quantity quota configuration of the root ZNode of the HBase service.

```
[zk:                :24002,                :24002,                :24002(CONNECTED) 7] listquota /hbase
absolute path is /zookeeper/quota/hbase
Output quota for /hbase count=1500000,bytes=10240
Output stat for /hbase count=59,bytes=1902
```

Step 4 Run the **getusage /hbase/splitWAL** command to check the ZNode quantity usage and check whether the ratio of **Node count** in the command output to the ZNode quantity quota is close to the alarm threshold.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **12007**, **19000**, or **19013** and the **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 6 Run the **getusage /hbase/replication** command to check the ZNode quantity usage and check whether the ratio of **Node count** in the command output to the ZNode quantity quota is close to the alarm threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **19006** and **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect the fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.191 ALM-19016 Quantity Quota Usage on ZooKeeper Exceeds the Threshold Severely

Alarm Description

The system checks the ZNode usage of the HBase service every 120 seconds. This alarm is generated when the znode usage of the HBase service exceeds the critical alarm threshold (90% by default).

This alarm is cleared when the quantity usage of the ZNode is less than the critical alarm threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19016	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Threshold	Specifies the threshold for generating the alarm.

Impact on the System

This alarm indicates that the quantity usage of the ZNode of HBase has exceeded the threshold severely. As a result, the write request of the HBase service fails.

Possible Causes

- DR is configured for HBase, and data synchronization fails or is slow in DR.
- A large number of WAL files are being split in the HBase cluster.

Handling Procedure

Check the quantity quota and usage of ZNodes.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, select the alarm whose ID is **19016**, and view the threshold in **Additional Information**.

Step 2 Log in to the HBase client as user **root**. Run the following command to go to the client installation directory:

cd *Client installation directory*

Run the following command to set environment variables:

source bigdata_env

If the cluster uses the security mode, run the following command to perform security authentication:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

kinit hbase

Enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 3 Run the **hbase zkcli** command to log in to the ZooKeeper client and run the **listquota /hbase** command to check the ZNode quantity quota of the HBase service. The ZNode root directory in the command is specified by the **zookeeper.znode.parent** parameter of the HBase service. The marked area in the following figure shows the quantity configuration of the root ZNode of the HBase service.



Step 4 Run the **getusage /hbase/splitWAL** command to check the ZNode usage and check whether the ratio of **Node count** in the command output to the znode quantity quota is close to the alarm threshold.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **12007**, **19000**, or **19013** and the **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 6 Run the **getusage /hbase/replication** command to check the ZNode usage and check whether the ratio of **Node count** in the command output to the ZNode quantity quota is close to the alarm threshold.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **19006** and **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect the fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.192 ALM-19017 Capacity Quota Usage on ZooKeeper Exceeds the Threshold

Alarm Description

The system checks the ZNode usage of the HBase service every 120 seconds. This alarm is generated when the system detects that the ZNodes capacity usage of the HBase service exceeds the alarm threshold (75% by default).

This alarm is cleared when the capacity usage of the ZNode capacity is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19017	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Threshold	Specifies the threshold for generating the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

This alarm indicates that the ZNodes capacity usage in the HBase service has exceeded the threshold. If this alarm is not handled in a timely manner, the problem severity may be escalated to **Critical**, affecting data writing.

Possible Causes

- DR is configured for HBase, and data synchronization fails or is slow in DR.
- A large number of WAL files are being split in the HBase cluster.

Handling Procedure

Check the capacity configuration and usage of ZNodes.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, select the alarm whose ID is **19017**, and view the threshold in **Additional Information**.

Step 2 Log in to the HBase client as user **root**. Run the following command to go to the client installation directory:

cd *Client installation directory*

Run the following command to set environment variables:

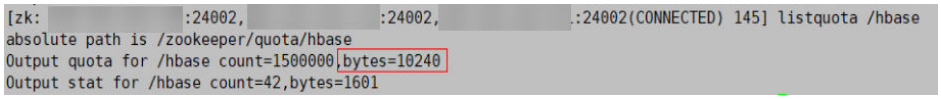
source bigdata_env

If the cluster uses the security mode, run the following command to perform security authentication:

kinit hbase

Enter the password as prompted.

Step 3 Run the **hbase zkcli** command to log in to the ZooKeeper client and run the **listquota /hbase** command to check the ZNode quantity quota of the HBase service. The ZNode root directory in the command is specified by the **zookeeper.znode.parent** parameter of the HBase service. The marked area in the following figure shows the quantity configuration of the root ZNode of the HBase service.



Step 4 Run the **getusage /hbase/splitWAL** command to check the capacity usage of the ZNode. Check whether the ratio of **Data size** to the ZNode capacity quota is close to the alarm threshold.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 On FusionInsight Manager, check whether the alarm whose ID is **12007**, **19000**, or **19013** and **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 7**.

Step 6 Run the **getusage /hbase/replication** command to check the capacity usage of the ZNode. Check whether the ratio of **Data size** to the ZNode capacity quota is close to the alarm threshold.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether the alarm whose ID is **19006** and **ServiceName** in **Location** is the current HBase service exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect the fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.193 ALM-19018 HBase Compaction Queue Size Exceeds the Threshold

Alarm Description

The system checks the HBase compaction queue size every 30 seconds. This alarm is generated when the compaction queue size exceeds the alarm threshold (**100** by default) for three consecutive times. This alarm is cleared when the compaction queue size is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19018	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Write pressure on the HBase node continues going up, and the disk I/O and CPU may be overloaded. Read and write requests are responded slowly or even time out.

Possible Causes

- The number of HBase RegionServers is too small.
- There are excessive regions on a single RegionServer of HBase.
- The HBase RegionServer heap size is small.
- Resources are insufficient.
- Related parameters are not configured properly.

Handling Procedure

Check whether related parameters are properly configured.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, check whether the alarm whose **Alarm ID** is **19008** or **19011** exists.

- If yes, click **View Help** next to the alarm and rectify the fault by referring to the help document. Then, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **HBase**. On the page that is displayed, click the **Configurations** tab then the **All Configurations** sub-tab, search for

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

hbase.hstore.compaction.min, **hbase.hstore.compaction.max**, **hbase.regionserver.thread.compaction.small**, and **hbase.regionserver.thread.compaction.throttle**, and set them to larger values.

Step 3 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect the fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.194 ALM-19019 Number of HBase HFiles to Be Synchronized Exceeds the Threshold

Alarm Description

The system checks the number of HFiles to be synchronized by the RegionServer of each HBase service instance every 30 seconds. This indicator can be viewed on the RegionServer role monitoring page. This alarm is generated when the number of HFiles to be synchronized on a RegionServer exceeds the threshold (exceeding 128 for 20 consecutive times by default). To change the threshold, choose **O&M** > **Alarm** > **Threshold Configuration** > *Name of the desired cluster* > **HBase** . This alarm is cleared when the number of HFiles to be synchronized is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19019	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

A large number of HFile files are stacked. Data is inconsistent between the active and standby nodes, and the latest data cannot be read from the standby cluster during an active/standby switchover or during HBase dual-read. If the fault persists, the storage space of the active cluster and ZooKeeper nodes will be used up. As a result, the active cluster service will be interrupted.

Possible Causes

- The network is abnormal.
- The RegionServer region distribution is unbalanced.
- The HBase service scale of the standby cluster is too small.

Handling Procedure

View alarm location information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **19019**, and view the service instance and host name in **Location**.

Check the network connection between RegionServers on active and standby clusters.

Step 2 Run the **ping** command to check whether the network connection between the faulty RegionServer node and the host where RegionServer of the standby cluster resides is normal.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 Contact the network administrator to restore the network.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 After the network recovers, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the RegionServer region distribution in the active cluster.

Step 5 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase**. Click **HMaster(Active)** to go to the web UI of the HBase instance and check whether regions are evenly distributed on the Region Server.

Step 6 Log in to the faulty RegionServer node as user **omm**.

Step 7 Run the following commands to go to the client installation directory and set the environment variable:

cd *Client installation directory*

source bigdata_env

If the cluster uses the security mode, perform security authentication. Run the **kinit hbase** command and enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 8 Run the following commands to check whether the load balancing function is enabled.

hbase shell

balancer_enabled

- If yes, go to **Step 10**.
- If no, go to **Step 9**.

Step 9 Run the following commands in HBase Shell to enable the load balancing function and check whether the function is enabled.

balance_switch true

balancer_enabled

Step 10 Run the **balancer** command to manually trigger the load balancing function.

NOTE

You are advised to enable and manually trigger the load balancing function during off-peak hours.

Step 11 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check the HBase service scale of the standby cluster.

Step 12 Expand the HBase cluster, add a node, and add a RegionServer instance on the node. Then, perform **Step 6** to **Step 10** to enable the load balancing function and manually trigger it.

Step 13 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase**. Click **HMaster(Active)** to go to the web UI of the HBase instance, refresh the page, and check whether regions are evenly distributed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Collect the fault information.

Step 15 On FusionInsight Manager of the standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 16 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 17 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 18 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.195 ALM-19020 Number of HBase WAL Files to Be Synchronized Exceeds the Threshold

Alarm Description

The system checks the number of WAL files to be synchronized by the RegionServer of each HBase service instance every 30 seconds. This indicator can be viewed on the RegionServer role monitoring page. This alarm is generated when the number of WAL files to be synchronized on a RegionServer exceeds the threshold (exceeding 128 for 20 consecutive times by default). To change the threshold, choose **O&M** > **Alarm** > **Threshold Configuration** > *Name of the desired cluster* > **HBase** . This alarm is cleared when the number of WAL files to be synchronized is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19020	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

A large number of WAL files are stacked. Data is inconsistent between the active and standby nodes, and the latest data cannot be read from the standby cluster during an active/standby switchover or during HBase dual-read. If the fault persists, the storage space of the active cluster and ZooKeeper nodes will be used up. As a result, the active cluster service will be interrupted.

Possible Causes

- The network is abnormal.
- The RegionServer region distribution is unbalanced.
- The HBase service scale of the standby cluster is too small.

Handling Procedure

View alarm location information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **19020**, and view the service instance and host name in **Location**.

Check the network connection between RegionServers on active and standby clusters.

Step 2 Run the **ping** command to check whether the network connection between the faulty RegionServer node and the host where RegionServer of the standby cluster resides is normal.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 Contact the network administrator to restore the network.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 After the network recovers, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the RegionServer region distribution in the active cluster.

Step 5 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase**. Click **HMaster(Active)** to go to the web UI of the HBase instance and check whether regions are evenly distributed on the Region Server.

Step 6 Log in to the faulty RegionServer node as user **omm**.

Step 7 Run the following commands to go to the client installation directory and set the environment variable:

cd *Client installation directory*

source bigdata_env

If the cluster uses the security mode, perform security authentication. Run the **kinit hbase** command and enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 8 Run the following commands to check whether the load balancing function is enabled.

hbase shell

balancer_enabled

- If yes, go to **Step 10**.
- If no, go to **Step 9**.

Step 9 Run the following commands in HBase Shell to enable the load balancing function and check whether the function is enabled.

balance_switch true

balancer_enabled

Step 10 Run the **balancer** command to manually trigger the load balancing function.

NOTE

You are advised to enable and manually trigger the load balancing function during off-peak hours.

Step 11 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check the HBase service scale of the standby cluster.

Step 12 Expand the HBase cluster, add a node, and add a RegionServer instance on the node. Then, perform **Step 6** to **Step 10** to enable the load balancing function and manually trigger it.

Step 13 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HBase**. Click **HMaster(Active)** to go to the web UI of the HBase instance, refresh the page, and check whether regions are evenly distributed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Collect the fault information.

Step 15 On FusionInsight Manager of the standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 16 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 17 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 18 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.196 ALM-19021 Handler Usage of RegionServer Exceeds the Threshold

Alarm Description

The system checks the RegionServer handler usage of each HBase service instance every 30 seconds. This alarm is generated when the handler usage of a RegionServer exceeds the threshold (90% for five consecutive times by default). This alarm is cleared if the handler usage is lower than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19021	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

RegionServer may not be able to provide services externally. As a result, the concurrent read/write processing become slow, or requests fail.

Possible Causes

- The value of a handler is too small.
- Hotspotting occurs.

Handling Procedure

View alarm location information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**, locate the row that contains the alarm whose **Alarm ID** is **19021**, and view the service instance and host name in **Location**.

Check the handler configuration.

Step 2 Choose **Cluster** > **Services** > **HBase** and click the **Configurations** tab. In the upper right corner of the page, search for **hbase.regionserver.handler.count** and check whether its value is too small. The default value is **200**.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Change the value of this parameter to a larger value and save the configuration. Choose **Cluster** > **Services** > **HBase**, click the **Instance** tab, select the affected RegionServer instances, and choose **More** > **Instance Rolling Restart**. In the displayed dialog box, enter the username and password. In the **Instance Rolling Restart** dialog box, click **OK** and wait until the rolling restart is complete.

Step 4 After the configuration takes effect, check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Check whether cluster hotspotting occurs.

Step 5 On FusionInsight Manager, choose **Cluster** > **Services** > **HBase** and click **HMaster(Active)** following **HMaster WebUI** to go to the web UI of the HBase instance. In the **Region Servers** area of the **Home** page, click **Requests** and check whether the requests in the **Filtered Read Request Count** and **Write Request Count** columns are evenly distributed.



- If yes, go to **Step 13**.
- If no, go to **Step 6**.

Step 6 Check whether regions are evenly distributed.

On FusionInsight Manager, choose **Cluster** > **Services** > **HBase** and click **HMaster(Active)** following **HMaster WebUI** to go to the web UI of the HBase instance. In the **Region Servers** area of the **Home** page, click **Base Stats** and check whether the regions in the **Num.Regions** column are evenly distributed.



- If yes, go to **Step 13**.
- If no, go to **Step 7**.

Step 7 Log in to the faulty RegionServer node as user **omm**.

Step 8 Run the following commands to go to the client installation directory and set the environment variable:

cd *Client installation directory*

source bigdata_env

If the cluster uses the security mode, run the following command to perform security authentication:

kinit hbase

Enter the password as prompted (obtain the password from the MRS cluster administrator).

Step 9 Run the following commands to check whether the load balancing function is enabled. If the command output is **true**, the load balancing function is enabled.

hbase shell

balancer_enabled

```
hbase:004:0> balancer_enabled
true
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
Took 0.0165 seconds
=> true
```

- If yes, go to **Step 13**.
- If no, go to **Step 10**.

Step 10 Run the following commands in HBase Shell to enable the load balancing function and check whether the function is enabled.

balance_switch true

balancer_enabled

NOTE

You are advised to enable and manually trigger the load balancing function during off-peak hours.

Step 11 Run the following command to manually trigger' the load balancing function:

balancer

Step 12 After the load balancing is complete, log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect the fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 15 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.197 ALM-19022 HBase Hotspot Detection Is Unavailable

Alarm Description

When the MetricController instance is installed for HBase, the alarm module checks the health status of the active HBase MetricController instance every 120

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

seconds. This alarm is generated when the active HBase MetricController instance does not exist or is unavailable and the hotspot detection function is unavailable.

This alarm is cleared when the active HBase MetricController instance recovers.

NOTE

This alarm applies only to MRS 3.3.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19022	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The HBase hotspot detection function is unavailable. Services are not affected. However, if request/data skew occurs, the system cannot report alarms and automatically recovers from hotspotting. Service requests may cause node overload, slow response, and request timeout.

Possible Causes

- The ZooKeeper service is abnormal.
- The HBase service is abnormal.
- In the current HBase service, the MetricController instance on the same node as the active HMaster instance is not started.
- The network is abnormal.

Handling Procedure

Check the ZooKeeper service status.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 In the service list on FusionInsight Manager, check whether **Running Status** of ZooKeeper is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 2**.

Step 2 In the alarm list, check whether **ALM-13000 ZooKeeper Service Unavailable** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by performing the operations provided for **ALM-13000 ZooKeeper Service Unavailable**.

Step 4 Wait for several minutes and check whether the alarm **HBase Hotspot Detection Is Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the HBase service status.

Step 5 In the service list on FusionInsight Manager, check whether **Running Status** of HBase is **Normal**.

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 In the alarm list, check whether the alarm ALM-19000 HBase Service Unavailable exists.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Rectify the fault by following the steps provided for **ALM-19000 HBase Service Unavailable**.

Step 8 Wait for several minutes and check whether the alarm **HBase Hotspot Detection Is Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check whether the MetricController instance deployed on the same node as the active HMaster instance is started.

Step 9 On FusionInsight Manager, choose **Cluster** > **Service** > **HBase**, and click **Instances** to check whether the **MetricController(Active)** instance exists.

- If yes, go to **Step 12**.
- If no, go to **Step 10**.

Step 10 Select the MetricController instance whose management IP address is the same as that of the active HMaster instance, and click **Start Instance**.

Step 11 After the MetricController instance is restarted, check whether the alarm **HBase Hotspot Detection Is Unavailable** is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 12**.

Check the network connectivity between the started MetricController instances and the active HMaster node.

Step 12 Log in to the node where the active HMaster instance is deployed and run **ping** *IP address of the node where the standby MetricController instance is deployed* to check whether the network connection between the started MetricController instances and the host where the active HMaster instance is deployed is normal.

- If yes, go to **Step 15**.
- If no, go to **Step 13**.

Step 13 Contact the network administrator to restore the network.

Step 14 After the network recovers, check whether the alarm **HBase Hotspot Detection Is Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Collect fault information.

Step 15 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 16 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 17 In the **Host** area, select the host where the HMaster instance is deployed.

Step 18 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 19 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.198 ALM-19023 Region Traffic Restriction for HBase

Alarm Description

When the MetricController instance is installed for the HBase service, self-healing from hotspotting is automatically enabled. The alarm module checks whether there are regions whose request traffic is restricted due to hotspot issues in HBase every 120 seconds. This alarm is generated when the region where hotspot traffic is restricted is detected in HBase.

This alarm is cleared when the region is no longer a hotspot.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm applies only to MRS 3.3.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19023	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the traffic of a hotspot region is restricted, the number of handlers for processing the requests in the region is limited. As a result, services requesting the region may slow down or retry upon failure.

Possible Causes

Too many requests are directed to a single region when the HBase service is accessed.

Handling Procedure

Check whether there are too many requests in a single region of HBase.

Step 1 Log in to FusionInsight Manager, and Choose **O&M** > **Alarm** > **Alarms**.

Step 2 In **Additional Information** of **Region Traffic Restriction for HBase**, view the reported table name and region information.

Step 3 On FusionInsight Manager, choose **Cluster** > **Service** > **HBase** and click the hyperlink on the right of HMaster web UI.

Step 4 Click **Table Details** and adjust service configurations in the region where the table in **Step 2** is deployed.

Step 5 Wait a moment and then check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 8 In the **Host** area, select the host where the HMaster instance is deployed.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm will be automatically cleared.

Related Information

None.

9.11.199 ALM-19024 RPC Requests P99 Latency on RegionServer Exceeds the Threshold

Alarm Description

The system checks P99 latency for RPC requests on each RegionServer instance of the HBase service every 30 seconds. This alarm is generated when P99 latency for RPC requests on a RegionServer exceeds the threshold for 10 consecutive times.

This alarm is cleared when P99 latency for RPC requests on a RegionServer instance is less than or equal to the threshold.

This alarm applies only to MRS 3.3.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19024	• **Critical**: The default threshold is 10 seconds. • **Major**: The default threshold is 5 seconds.	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If RPC requests P99 latency exceeds the threshold, the RegionServer cannot deliver normal service performance externally. For latency-sensitive services, a large number of service read and write requests may time out.

Possible Causes

- RegionServer GC duration is too long.
- The HDFS RPC response is too slow.
- RegionServer request concurrency is too high.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **19024**, and view the service instance and host name in **Location**.

Check the GC duration of RegionServer.

Step 2 In the alarm list on FusionInsight Manager, check whether the "HBase GC Duration Exceeds the Threshold" alarm is generated for the service instance in **Step 1**.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Rectify the fault by following the handling procedure of **ALM-19007 HBase GC Duration Exceeds the Threshold**.

Step 4 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check HDFS RPC response time.

Step 5 In the alarm list on FusionInsight Manager, check whether alarm "Average NameNode RPC Processing Time Exceeds the Threshold" is generated for the HDFS service on which the HBase service depends.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Rectify the fault by following the handling procedure of **ALM-14021 Average NameNode RPC Processing Time Exceeds the Threshold**.

Step 7 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check the number of concurrent processes on a RegionServer.

Step 8 In the alarm list on FusionInsight Manager, check whether the "Handler Usage of RegionServer Exceeds the Threshold" alarm is generated for the service instance in **Step 1**.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Rectify the fault by following the handling procedure of **ALM-19021 Handler Usage of RegionServer Exceeds the Threshold**.

Step 10 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 12 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 13 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.200 ALM-19025 Damaged StoreFile in HBase

Alarm Description

The system checks the **hdfs://hacluster/hbase/autocorrupt** and **hdfs://hacluster/hbase/MasterData/autocorrupt** directories on HDFS of each HBase service every 120 seconds. This alarm is generated when there are files in the directories.

This alarm is cleared when the **hdfs://hacluster/hbase/autocorrupt** and **hdfs://hacluster/hbase/MasterData/autocorrupt** directories do not exist or are empty.

This alarm applies only to MRS 3.3.0 or later.

NOTE

hdfs://hacluster indicates the name of the file system used by HBase, and **/hbase** indicates the root directory of HBase in the file system. You can log in to FusionInsight Manager, choose **Cluster** > **Services** > **HBase** and click **Configuration**. Search for **fs.defaultFS** and **hbase.data.rootdir**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19025	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Data in the file may be lost, and data queried by the service may be inconsistent.

Possible Causes

The StoreFile files are damaged.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **19025**, and view the service in **Location**.

Step 2 Log in to the node where the HDFS and HBase clients are installed as the client installation user and run the following commands:

cd *Client installation directory*

source bigdata_env

kinit *Component service user* (If Kerberos authentication is disabled for the cluster (the cluster is in normal mode), skip this step.)

Step 3 Check the damaged StoreFile file.

- Run the following command to check whether the **/hbase/autocorrupt** directory of HDFS is empty. If it is not, go to **Step 4**.

  hdfs dfs -ls -R hdfs://hacluster/hbase/autocorrupt

- Run the following command to check whether the **/hbase/MasterData/autocorrupt** directory of HDFS is empty. If it is not, go to **Step 9**.

  hdfs dfs -ls -R hdfs://hacluster/hbase/MasterData/autocorrupt

Step 4 Run the following command to restore the StoreFile files in the **hdfs://hacluster/hbase/autocorrupt** directory:

hdfs debug recoverLease -path hdfs://hacluster/hbase/autocorrupt/*Name space*/Table/Region/Column family/*StoreFile files*

Step 5 Check whether the damaged StoreFile files are restored. If the following information is displayed, the restoration is successful:

```
recoverLease SUCCEEDED on hdfs://hacluster/hbase/autocorrupt/
default/h1/865665fe32db62dadada68b644359809/cf1/95f210f931ad44c99e4028470be7d292
```

If yes, go to **Step 6**.

If no, go to **Step 9**.

Step 6 Run the following command to move the files back to the **hdfs://hacluster/hbase/data** directory:

hdfs dfs -mv hdfs://hacluster/hbase/autocorrupt/*Name space*/Table/Region/Column family/*StoreFile files***hdfs://hacluster/hbase/data/***Name space*/Table/Region/Column family/*StoreFile files*

Step 7 Run the following command on HBase Shell to bring the region online again:

hbase shell

unassign'*Region*'

assign'*Region*'

Step 8 Wait several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.201 ALM-19026 Damaged WAL Files in HBase

Alarm Description

The system checks the **hdfs://hacluster/hbase/corrupt** directory on the HDFS of each HBase service every 120 seconds. This alarm is generated when there are WAL files in the **/hbase/corrupt** directory.

This alarm is cleared when the **/hbase/corrupt** directory does not exist or does not contain WAL files.

This alarm applies only to MRS 3.3.0 or later.

NOTE

hdfs://hacluster indicates the name of the file system used by HBase, and **/hbase** indicates the root directory of HBase in the file system. You can log in to FusionInsight Manager, choose **Cluster** > **Services** > **HBase** and click **Configuration**. Search for **fs.defaultFS** and **hbase.data.rootdir**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
19026	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the data in the damaged file is not flushed to disks, the data will be lost. As a result, some data queried by the service is inconsistent.

Possible Causes

The WAL files are damaged.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **19026**, and view the service in **Location**.

Step 2 Log in to the node where the HDFS clients are installed as the client installation user and run the following commands:

cd *Client installation directory*

source bigdata_env

kinit *Component service user* (If Kerberos authentication is disabled for the cluster (the cluster is in normal mode), skip this step.)

Step 3 Run the following command to check the damaged WAL files and go to **Step 4**:

hdfs dfs -ls hdfs://hacluster/hbase/corrupt/*%2C*

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **HBase** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.202 ALM-20002 Hue Service Unavailable

Description

The system checks the Hue service status every 60 seconds. This alarm is generated if the Hue service is unavailable.

This alarm is cleared when the Hue service is normal.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
20002	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Users cannot perform interactive analysis and data processing with MRS on the Hue UI.

Possible Causes

- The internal KrbServer service on which the Hue service depends is abnormal.
- The internal DBService service on which the Hue service depends is abnormal.
- The network connection to the DBService is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check whether the KrbServer is abnormal.

Step 1 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services**. In the service list, check whether the **KrbServer** running status is **Normal**.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Restart the KrbServer service.

Step 3 Wait for several minutes. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the DBService is abnormal.

Step 4 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services**.

Step 5 In the service list, check whether the **DBService** running status is **Normal**.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Restart the DBService.

NOTICE

When the service is rebooted, it becomes unavailable and can disrupt business operations.

Step 7 Wait for several minutes. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the network connection to the DBService is normal.

Step 8 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Hue** > **Instance**, record the IP address of the active Hue.

Step 9 Log in to the active Hue.

Step 10 Run the **ping** command to check whether communication between the host that runs the active Hue and the hosts that run the DBService is normal. (Obtain the IP addresses of the hosts that run the DBService in the same way as that for obtaining the IP address of the active Hue.)

- If yes, go to **Step 13**.
- If no, go to **Step 11**.

Step 11 Contact the administrator to restore the network.

Step 12 Wait for several minutes. Check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 14 Select the following nodes in the required cluster from the **Service** drop-down list:

- Hue
- Controller

Step 15 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.203 ALM-23001 Loader Service Unavailable

Description

The system checks the Loader service availability every 60 seconds. This alarm is generated when the system detects that the Loader service is unavailable. This alarm is cleared when the Loader service is available.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
23001	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

When the Loader service is unavailable, the data loading, import, and conversion functions are unavailable.

Possible Causes

- The internal service on which the Loader service depends is abnormal.
  - The ZooKeeper service is abnormal.
  - The HDFS service is abnormal.
  - The DBService service is abnormal.
  - The Yarn service is abnormal.
  - MapReduce is abnormal.
- Environment fault: The network is abnormal, which the Loader service cannot communicate with the depended internal services and cannot provide services.
- Software fault: The Loader service cannot run properly.

Procedure

Check the ZooKeeper service status.

Step 1 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **ZooKeeper** to check whether the ZooKeeper running status is **Normal**.

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Choose **More** > **Restart Service** to restart the ZooKeeper service. In the alarm list, check whether **LoaderService Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 On the FusionInsight Manager, check whether the alarm list contains **Process Fault**.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 In the **Location** area of **Process Fault**, check whether **ServiceName** is **ZooKeeper**.

- If yes, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 7**.

Step 5 Rectify the fault by following the steps provided in **ALM-12007 Process Fault**.

Step 6 In the alarm list, check whether **Loader Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check the HDFS service status.

Step 7 On the FusionInsight Manager, check whether the alarm list contains **HDFS Service Unavailable**.

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

Step 8 Rectify the fault by following the steps provided in **ALM-14000 HDFS Service Unavailable**.

Step 9 In the alarm list, check whether **Loader Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check the DBService status.

Step 10 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService** to check whether the DBService running status is **Normal**.

- If yes, go to **Step 12**.
- If no, go to **Step 11**.

Step 11 Choose **More** > **Restart Service** to restart the DBService service. In the alarm list, check whether **LoaderService Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check the MapReduce service status.

Step 12 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **MapReduce** to check whether the MapReduce running status is **Normal**.

- If yes, go to **Step 16**.
- If no, go to **Step 13**.

Step 13 Choose **More** > **Restart Service** to restart the MapReduce service. In the alarm list, check whether **LoaderService Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Check the Yarn status.

Step 14 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** to check whether the Yarn running status is **Normal**.

- If yes, go to **Step 16**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 15**.

Step 15 Choose **More** > **Restart Service** to restart the Yarn service. In the alarm list, check whether **LoaderService Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Step 16 On the FusionInsight Manager, check whether the alarm list contains **Yarn Service Unavailable**.

- If yes, go to **Step 17**.
- If no, go to **Step 19**.

Step 17 Rectify the fault by following the steps provided in **ALM-18000 Yarn Service Unavailable**.

Step 18 In the alarm list, check whether **Loader Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 19**.

Check the network connection between Loader and dependent components.

Step 19 On the FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader**.

Step 20 Click **Instance** and the LoaderServer instance list is displayed.

Step 21 Record the **Management IP Address** in the row of **LoaderServer(Active)**.

Step 22 Log in to the host where the active LoaderServer runs as **omm** user using the IP address obtained in **Step 21**.

Step 23 Run the **ping** command to check whether communication between the host that runs the active LoaderServer and the hosts that run the dependent components. (The dependent components include ZooKeeper, DBService, HDFS, MapReduce and Yarn. Obtain the IP addresses of the hosts that run these services in the same way as that for obtaining the IP address of the active LoaderServer.)

- If yes, go to **Step 26**.
- If no, go to **Step 24**.

Step 24 Contact the administrator to restore the network.

Step 25 In the alarm list, check whether **Loader Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 26**.

Collect fault information.

Step 26 On the FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 27 Select the following nodes in the required cluster from the **Service** drop-down list:

- Zookeeper
- HDFS
- DBService

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Yarn
- MapReduce
- Loader

Step 28 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 29 On the FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader**.

Step 30 Choose **More** > **Restart Service**, and click **OK**.

Step 31 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 32**.

Step 32 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.204 ALM-23003 Loader Task Execution Failure

Description

This alarm is generated immediately when the system detects that the Loader job fails. This alarm is cleared when the failed job is manually handled by a user. This alarm must be manually cleared.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
23003	Minor	No

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
JobID	Specifies the ID of failed Loader job.
JobName	Specifies the failed Loader job.
UserName	Specifies the name of the user who submits the Loader job.
Details	Supplementary information for which the alarm is generated.

Impact on the System

This is a job-level alarm for Loader. The job execution fails, and you need to view specific logs to locate the failure cause. No execution result is returned. After the fault is rectified, you need to execute the task again. No impact on the Loader service.

Possible Causes

- Task parameters are incorrectly configured.
- Exceptions occur when Yarn is executing a job.

Procedure

Check whether task parameters are incorrectly configured.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and click the alarm drop-down list from the alarm list, obtain the **Alarm Cause**.

Step 2 If the alarm cause is "Failure to submit job", view error details in **Additional Information**, and go to the Loader WebUI to view the execution history of the job.

NOTE

By default, the **admin** user does not have the permissions to manage other components. If the page cannot be opened or the displayed content is incomplete when you access the native UI of a component due to insufficient permissions, you can manually create a user with the permissions to manage that component.

Step 3 Submit the task again.

Step 4 Check whether the task executed successfully.

- If yes, go to **Step 9**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Check whether exceptions occur when Yarn is executing a job.

Step 5 On FusionInsight Manager, click the alarm drop-down list from the alarm list, obtain the **Alarm Cause**.

Step 6 Check whether the Yarn activity is executed properly in the **Alarm Cause**. If the alarm cause is "Yarn execution failed", the Yarn activity is abnormal.

- If yes, go to **Step 7**.
- If no, go to **Step 10**.

Step 7 Submit the task again.

Step 8 Please check whether the task executed successfully.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 In the alarm list, click **Clear** from **Operation** to manually clear the alarm. No further action is required.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 11 Select the following nodes in the required cluster from the **Service** drop-down list:

- DBService
- HDFS
- Loader
- MapReduce
- Yarn
- ZooKeeper

Step 12 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system does not automatically clear this alarm, and you need to manually clear the alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.205 ALM-23004 Loader Heap Memory Usage Exceeds the Threshold

Description

The system checks the heap memory usage of the Loader service every 60 seconds. The alarm is generated when the heap memory usage of a Loader instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times. The alarm is cleared when the heap memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
23004	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Full GC occurs frequently in Loader. The performance deteriorates and page responses are slow. If the memory overflows, Loader may fail to provide services for external systems. The Loader page cannot be accessed, interfaces cannot be called, and active/standby switchover is frequently performed due to exceptions.

Possible Causes

The heap memory of the Loader instance is overused or the heap memory is inappropriately allocated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Loader Heap Memory Usage Exceeds the Threshold** > **Location.** Check the host name of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **Memory** > **Loader Heap Memory Resource Percentage**. Click **OK**.

Step 3 Check whether the used heap memory of Loader reaches the threshold (the default value is 95% of the maximum heap memory) specified for Loader.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Configurations**. Click **All Configurations**. Increase the value of **-Xmx** in **GC_OPTS** as required, and click **Save**. Click **OK**.

NOTE

- If this alarm is generated, the heap memory configured for the current Loader instance is insufficient for data transmission. You are advised to open the instance monitoring page, display the Loader heap memory resource status monitoring chart, and observe the change trend of the heap memory used by Loader in the monitoring chart. Then change the value of **-Xmx** to twice the current heap memory usage or to another value to meet site requirements.
- When setting the heap memory, you can set **-Xms** and **-Xmx** to similar values to avoid performance deterioration caused by heap size adjustment after each GC.
- Ensure that the sum of **-Xmx** and **XX:MaxPermSize** is not greater than the physical memory of the node server.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTE

During service or instance restart, Loader cannot provide services to external systems. New jobs cannot be submitted, but jobs being executed are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Loader** in the required cluster from the **Service** drop-down list.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected fault logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.206 ALM-23005 Loader Non-Heap Memory Usage Exceeds the Threshold

Description

The system checks the non-heap memory usage of the Loader service every 30 seconds. The alarm is generated when the non-heap memory usage of a Loader instance exceeds the threshold (80% of the maximum memory) for 5 consecutive times. The alarm is cleared when the non-heap memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
23005	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The Loader page may fail to be accessed and cannot provide services for external systems.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The non-heap memory of the Loader instance is overused or the non-heap memory is inappropriately allocated.

Procedure

Check non-heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Loader Non-Heap Memory Usage Exceeds the Threshold** > **Location.** Check the host name of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **Memory** > **Loader Non Heap Memory Resource Percentage**. Click **OK**.

Step 3 Check whether the used non-heap memory of Loader reaches the threshold (the default value is 80% of the maximum non-heap memory) specified for Loader.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Configurations**. Click **All Configurations** Search **LOADER_GC_OPTS** in the search box. If the **-XX: MaxPermSize** parameter is not configured, set the initial value to **-XX: MaxPermSize=256M** for the first time. (If the alarm persists after the first adjustment, perform the second adjustment by referring to the following note.) And click **Save**. Click **OK**.

NOTE

If this alarm is generated, the non-heap memory configured for the current Loader instance is insufficient for the service scenario. You are advised to open the instance monitoring page, open the Loader non-heap memory resource status monitoring chart, and observe the change trend of the non-heap memory used by Loader in the monitoring chart. Then change the value of **-XX:MaxPermSize** to twice the current non-heap memory usage or to another value to meet site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTE

During service or instance restart, Loader cannot provide services to external systems. New jobs cannot be submitted, but jobs being executed are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Loader** in the required cluster from the **Service** drop-down list.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.207 ALM-23006 Loader Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the Loader service every 30 seconds. The alarm is generated when the direct memory usage of a Loader instance exceeds the threshold (80% of the maximum memory) for 5 consecutive times. The alarm is cleared when the direct memory usage of Loader is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
23006	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Loader may fail to provide services for external systems. I/O or socket exceptions occur, and active/standby switchovers occur frequently.

Possible Causes

The direct memory of the Loader instance is overused or the direct memory is inappropriately allocated.

Procedure

Check direct memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Loader Direct Memory Usage Exceeds the Threshold** > **Location**. Check the host name of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **Memory** > **Loader Memory Usage Statistics**. Click **OK**.

Step 3 Check whether the used direct memory of Loader reaches the threshold (the default value is 80% of the maximum direct memory) specified for Loader.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Configurations**. Click **All Configurations**. Search **LOADER_GC_OPTS** in the search box. Increase the value of **-XX:MaxDirectMemorySize** as required, and click **Save**. Click **OK**.

NOTE

If this alarm is generated, the direct memory configured for the current Loader instance is insufficient for the service scenario. You are advised to open the instance monitoring page, display the Loader direct memory resource status monitoring chart, and observe the change trend of the direct memory used by Loader in the monitoring chart. Then change the value of **-XX:MaxDirectMemorySize** to twice the current direct memory usage or to another value to meet site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTE

During service or instance restart, Loader cannot provide services to external systems. New jobs cannot be submitted, but jobs being executed are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Loader** in the required cluster from the **Service** drop-down list.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.208 ALM-23007 Garbage Collection (GC) Time of the Loader Process Exceeds the Threshold

Description

The system checks GC time of the Loader process every 60 seconds. The alarm is generated when GC time of the Loader process exceeds the threshold (default value: **12 seconds**) for 5 consecutive times. The alarm is cleared when GC time is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
23007	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Full GC occurs frequently, and the Loader service responds slowly. The Loader service may even break down and cannot provide services properly.

Possible Causes

Full GC occurs frequently, and the Loader service responds slowly. The Loader service may even break down and cannot provide services properly.

Procedure

Check GC time.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Garbage Collection (GC) Time of the Loader Process Exceeds the Threshold** > **Location**. Check the host name of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **GC** > **Garbage Collection (GC) Time of Loader**. Click **OK**.

Step 3 Check whether GC time of the Loader process every second exceeds the threshold (default value: **12 seconds**).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Loader** > **Configurations**. Click **All Configurations**. Search **LOADER_GC_OPTS** in the search box. Increase the value of **-Xmx** as required, and click **Save**. Click **OK**.

NOTE

If this alarm is generated, the heap memory configured for the current Loader instance cannot meet the heap memory required for data transmission. You are advised to handle the problem by referring to **Step 4** in section **ALM-23004 Loader Heap Memory Usage Exceeds the Threshold**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

During service or instance restart, Loader cannot provide services to external systems. New jobs cannot be submitted, but jobs being executed are not affected.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Loader** in the required cluster from the **Service** drop-down list.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.209 ALM-24000 Flume Service Unavailable

Alarm Description

The alarm module checks the Flume service status every 180 seconds. This alarm is generated if the Flume service is abnormal.

This alarm is automatically cleared after the Flume service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24000	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Flume cannot work and data transmission is interrupted.

Possible Causes

All Flume instances are faulty.

Handling Procedure

Step 1 Log in to a Flume node as user **omm** and run the **ps -ef|grep "flume.role=server"** command to check whether the Flume process exists on the node.

- If yes, go to **Step 3**.
- If no, restart the faulty Flume node or Flume service and go to **Step 2**.

Step 2 In the alarm list, check whether alarm "Flume Service Unavailable" is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Collect the fault information.

Step 3 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 4 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 5 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 6 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.210 ALM-24001 Flume Agent Exception

Alarm Description

The Flume agent monitoring module monitors the Flume agent status. This alarm is generated when the Flume agent process is faulty (checked every 5 seconds) or the Flume agent fails to start (an alarm is reported immediately).

This alarm is cleared when the Flume agent process recovers, Flume agent starts successfully, and the alarm handling is completed.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24001	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
AgentId	Specifies the ID of the agent for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The Flume agent instance for which the alarm is generated cannot provide services properly, and the data transmission tasks of the instance are temporarily interrupted. Real-time data is lost during real-time data transmission.

Possible Causes

- The **JAVA_HOME** directory does not exist, or the Java permission is incorrect.
- The Flume agent directory permission is incorrect.
- The Flume agent fails to start.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check whether the JAVA_HOME directory exists or whether the JAVA permission is correct.

Step 1 Log in to the host for which the alarm is generated as user **root**.

Step 2 Obtain the installation directory of the Flume client for which the alarm is generated. (The value of **AgentId** can be obtained from **Location** of the alarm.)

ps -ef|grep *AgentId* **| grep -v grep | awk -F 'conf-file ' '{print $2}' | awk -F 'fusioninsight' '{print $1}'**

Step 3 Run the **su -** *Flume installation user* command to switch to the Flume installation user and run the **cd** *Flume client installation directory*/**fusioninsight-flume-1.9.0/conf/** command to go to the Flume configuration directory.

Step 4 Run the **cat ENV_VARS | grep JAVA_HOME** command.

Step 5 Check whether the **JAVA_HOME** directory exists. If both the command output in **Step 4** and **ll $JAVA_HOME/** are not empty, the **JAVA_HOME** directory exists.

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

Step 6 Specify a correct **JAVA_HOME** directory.

Step 7 Run the **$JAVA_HOME/bin/java -version** command to check whether the Flume agent running user has the Java execution permission. If the Java version is displayed in the command output, the Java permission meets the requirement. Otherwise, the Java permission does not meet the requirement.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

NOTE

JAVA_HOME is the environment variable exported during Flume client installation. You can also go to *Flume client installation directory*/**fusioninsight-flume-1.9.0/conf** and run the **cat ENV_VARS | grep JAVA_HOME** command to view the variable value.

Step 8 Run the **chmod 750 $JAVA_HOME/bin/java** command to grant the Java execution permission to the Flume agent running user.

Check the directory permission of the Flume agent.

Step 9 Log in to the host for which the alarm is generated as user **root**.

Step 10 Run the following command to switch to the Flume agent installation directory:

cd *Flume client installation directory*/**fusioninsight-flume-1.9.0/conf/**

Step 11 Run the **ls -al * -R** command to check whether any file owner is the user running the Flume agent.

- If yes, go to **Step 12**.
- If no, run the **chown** command to change the file owner to the user who runs the Flume agent.

Check the Flume agent configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Run the **cat properties.properties | grep spooldir** and **cat properties.properties | grep TAILDIR** commands to check whether the Flume source type is spoolDir or tailDir. If any command output is displayed, the Flume source type is spoolDir or tailDir.

- If yes, go to **Step 13**.
- If no, go to **Step 17**.

Step 13 Check whether the data monitoring directory exists.

- If yes, go to **Step 15**.
- If no, go to **Step 14**.

NOTE

Run the **cat properties.properties | grep spoolDir** command to view the spoolDir monitoring directory.

```
[root@[redacted] /tmp/test/fusioninsight-flume-1.9.0/conf]# cat properties.properties | grep spoolDir
client.sources.aa1.spoolDir = /opt/liuxingcheng/flumeclient/sourcedata/flumesourcedata1
[root@[redacted] /tmp/test/fusioninsight-flume-1.9.0/conf]#
```

Run the **cat properties.properties | grep parentDir** command to view the tailDir monitoring directory.

```
[root@[redacted] /tmp/test/fusioninsight-flume-1.9.0/conf]# cat properties.properties | grep parentDir
server.sources.AAAA.filegroups.F1.parentDir = /tmp/flumetest/taildir_data
[root@[redacted] /tmp/test/fusioninsight-flume-1.9.0/conf]#
```

Step 14 Specify a correct data monitoring directory.

Step 15 Check whether the Flume agent user has the read, write, and execute permissions on the monitoring directory specified in **Step 13**.

- If yes, go to **Step 17**.
- If no, go to **Step 16**.

NOTE

Go to the monitoring directory as the Flume running user. If files can be created, the Flume running user has the read, write, and execute permissions on the monitoring directory.

Step 16 Run the **chmod 777** *Flume monitoring directory* command to grant the Flume agent running user the read, write, and execute permissions on the monitoring directory specified in **Step 13**.

Step 17 Check whether the components connected to the Flume sink are in safe mode.

- If yes, go to **Step 18**.
- If no, go to **Step 23**.

NOTE

If the sinks in the **properties.properties** configuration file are the HDFS sink and HBase sink, and the configuration file contains a keytab file, the components connected to the Flume sink are in safe mode.

If the sink in the **properties.properties** configuration file is the Kafka sink and ***.security.protocol** is set to **SASL_PLAINTEXT** or **SASL_SSL**, Kafka connected to the Flume sink is in safe mode.

Step 18 Run the **ll** *ketab path* command to check whether the keytab authentication path specified by the ***.kerberosKeytab** parameter in the configuration file exists.

- If yes, go to **Step 20**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 19**.

  NOTE

  To view the ketab path, run the **cat properties.properties | grep keytab** command.

Step 19 Change the value of **kerberosKeytab** in **Step 18** to the custom keytab path and go to **Step 21**.

Step 20 Go to **Step 18** and check whether the Flume agent running user has the permission to access the keytab authentication file. If the keytab path is returned, the user has the permission. Otherwise, the user does not have the permission.

- If yes, go to **Step 22**.
- If no, go to **Step 21**.

Step 21 Run the **chmod 755** *ketab file* command to grant the read permission on the keytab file specified in **Step 19**, and restart the Flume process.

Step 22 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 23**.

Collect fault information.

Step 23 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 24 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 25 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 26 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.211 ALM-24003 Flume Client Connection Interrupted

Alarm Description

The alarm module monitors the port connection status on the Flume server. This alarm is generated if the Flume server fails to receive a connection message from the Flume client in three consecutive minutes.

This alarm is cleared after the Flume server receives a connection message from the Flume client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24003	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
Client IP Address	Specifies the IP address of the Flume client.
Client Name	Specifies the agent name of the Flume client.
Sink Name	Specifies the sink name of Flume Agent.

Impact on the System

The communication between the Flume client and the server fails. The Flume client cannot send data to the Flume server.

Possible Causes

- The network connection between the Flume client and the server is faulty.
- The Flume client's process is abnormal.
- The Flume client is incorrectly configured.

Handling Procedure

Check the network connection between the Flume client and the server.

Step 1 Log in to the host whose IP address is specified by **Flume ClientIP** in the alarm information as user **root**.

Step 2 Run the **ping** *Flume server IP address* command to check whether the network connection between the Flume client and the server is normal.

- If yes, go to **Step 3**.
- If no, go to **Step 11**.

Check whether the Flume client's process is normal.

Step 3 Log in to the host whose IP address is specified by **Flume ClientIP** in the alarm information as user **root**.

Step 4 Run the **ps -ef|grep flume |grep client** command to check whether the Flume client process exists.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 5**.
- If no, go to **Step 11**.

Check the Flume client configuration.

Step 5 Log in to the host whose IP address is specified by **Flume ClientIP** in the alarm information as user **root**.

Step 6 Run the **cd** *Flume client installation directory*/**fusioninsight-flume-1.9.0/conf/** command to go to Flume's configuration directory.

Step 7 Run the **cat properties.properties** command to query the current configuration file of the Flume client.

Step 8 Check whether the **properties.properties** file is correctly configured according to the configuration description of the Flume agent.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Modify the **properties.properties** configuration file.

Check whether the alarm is cleared.

Step 10 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect the fault information.

Step 11 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 12 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 13 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Collect logs in the **/var/log/Bigdata/flume-client** directory on the Flume client using a transmission tool.

Step 15 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.212 ALM-24004 Exception Occurs When Flume Reads Data

Alarm Description

The alarm module monitors the Flume source status. This alarm is generated when the duration in which the source cannot read data exceeds the threshold.

The default threshold is **0**, indicating that this function is disabled. You can change the threshold by modifying the **properties.properties** file in the **conf** directory. Specifically, modify the **NoDatatime** parameter of required the source.

The alarm is cleared when the source reads the data and the alarm handling is complete.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24004	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
AgentId	Specifies the ID of the agent for which the alarm was generated.
ComponentType	Specifies the type of the component for which the alarm was generated.
ComponentName	Specifies the name of the component for which the alarm was generated.

Impact on the System

If data is found in the data source but the Flume source continuously fails to read data, the collection is stopped.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The Flume source is faulty, so data cannot be sent.
- The network is faulty, so the data cannot be sent.

Handling Procedure

Check whether the Flume source is faulty.

Step 1 Open the **properties.properties** configuration file on the local PC, search for keyword **type = spooldir** in the file, and check whether the Flume source type is spoolDir.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 View the spoolDir monitoring directory to check whether all files are already transferred.

- If yes, no further action is required.
- If no, go to **Step 5**.

NOTE

The monitoring directory of spoolDir is specified by the **.spoolDir** parameter in the **properties.properties** configuration file. If all files in the monitoring directory have been transferred, the file name extension of all files in the monitoring directory is **.COMPLETED**.

Step 3 Open the **properties.properties** configuration file on the local PC, search for **org.apache.flume.source.kafka.KafkaSource** in the file, and check whether the Flume source type is Kafka.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 Check whether the topic data configured by the Kafka source has been used up.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 On FusionInsight Manager, choose **Cluster**, click the name of the desired cluster, and choose **Services** > **Flume** > **Instances**.

Step 6 Go to the Flume instance page of the faulty node to check whether the **Source Speed Metrics** in the alarm is **0**.

- If yes, go to **Step 11**.
- If no, go to **Step 7**.

Check the network connectivity between the node with the IP address configured for the Flume source and the faulty node.

Step 7 Open the **properties.properties** configuration file on the local PC, search for **type = avro** in the file, and check whether the Flume source type is Avro.

- If yes, go to **Step 8**.
- If no, go to **Step 11**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Log in to the faulty node as user **root**, and run the **ping** *IP address of the Flume source* command to check whether the peer host can be pinged successfully.

- If yes, go to **Step 11**.
- If no, go to **Step 9**.

Step 9 Contact the network administrator to restore the network.

Step 10 Wait for a while and check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 12 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 13 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.213 ALM-24005 Exception Occurs When Flume Transmits Data

Alarm Description

The alarm module monitors the capacity status of Flume Channel. The alarm is generated immediately when the duration that Channel is fully occupied exceeds the threshold or the number of times that Source fails to send data to Channel exceeds the threshold.

The default threshold is **10**. You can change the threshold by modifying the **channelfullcount** parameter of the related channel in the **properties.properties** configuration file in the **conf** directory.

The alarm is cleared when the space of Flume Channel is released and the alarm handling is complete.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24005	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
AgentId	Specifies the ID of the agent for which the alarm was generated.
ComponentType	Specifies the type of the component for which the alarm was generated.
ComponentName	Specifies the name of the component for which the alarm was generated.

Impact on the System

If the disk usage of Flume Channel increases continuously, the time required for importing data to a specified destination prolongs. When the disk usage of Flume Channel reaches 100%, the Flume agent process pauses.

Possible Causes

- Flume Sink is faulty, so the data cannot be sent.
- The network is faulty, so the data cannot be sent.

Handling Procedure

Check whether Flume Sink is faulty.

Step 1 Open the **properties.properties** configuration file on the local PC, search for **type = hdfs** in the file, and check whether the Flume sink type is HDFS.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 On FusionInsight Manager, check whether **HDFS Service Unavailable** alarm is generated in the alarm list and whether the HDFS service is stopped in the service list.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the alarm is reported, clear it according to the handling suggestions of **ALM-14000 HDFS Service Unavailable**; if the HDFS service is stopped, start it. Then, go to **Step 7**.
- If no, go to **Step 7**.

Step 3 Open the **properties.properties** configuration file on the local PC, search for **type = hbase** in the file, and check whether the Flume sink type is HBase.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 On FusionInsight Manager, check whether **HBase Service Unavailable** alarm is generated in the alarm list and whether the HBase service is stopped in the service list.

- If the alarm is reported, clear it according to the handling suggestions of **ALM-19000 HBase Service Unavailable**; if the HBase service is stopped, start it. Then, go to **Step 7**.
- If no, go to **Step 7**.

Step 5 Open the **properties.properties** configuration file on the local PC, search for **org.apache.flume.sink.kafka.KafkaSink** in the file, and check whether the Flume sink type is Kafka.

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 On FusionInsight Manager, check whether **Kafka Service Unavailable** alarm is generated in the alarm list and whether the Kafka service is stopped in the service list.

- If the alarm is reported, clear it according to the handling suggestions of **ALM-38000 Kafka Service Unavailable**; if the Kafka service is stopped, start it. Then, go to **Step 7**.
- If no, go to **Step 7**.

Step 7 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Flume** > **Instance**.

Step 8 Go to the Flume instance page of the faulty node to check whether the indicator **Sink Speed Metrics** is 0.

- If yes, go to **Step 13**.
- If no, go to **Step 9**.

Check the network connection between the faulty node and the node that corresponds to the Flume Sink IP address.

Step 9 Open the **properties.properties** configuration file on the local PC, search for **type = avro** in the file, and check whether the Flume sink type is Avro.

- If yes, go to **Step 10**.
- If no, go to **Step 13**.

Step 10 Log in to the faulty node as user **root**, and run the **ping** *IP address of the Flume sink* command to check whether the peer host can be pinged successfully.

- If yes, go to **Step 13**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 11**.

Step 11 Contact the network administrator to restore the network.

Step 12 In the alarm list, check whether the alarm is cleared after a period.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect the fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.214 ALM-24006 Heap Memory Usage of Flume Server Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the Flume service every 60 seconds. This alarm is generated when the heap memory usage of the Flume instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times. This alarm is cleared when the heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24006	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the heap memory overflows, the service may break down and the Flume instance may be unavailable.

Possible Causes

The heap memory of the Flume instance is overused or the heap memory is inappropriately allocated.

Handling Procedure

Check the heap memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **Flume Heap Memory Usage Exceeds the Threshold**, and view the **Location** information. Check the name of the host for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Flume**. On the page that is displayed, click the **Instance** tab. On the displayed tab page, select the role corresponding to the host name for which the alarm is generated and select **Customize** from the drop-down list in the upper right corner of the chart area. Choose **Agent** and select **Flume Heap Memory Resource Percentage**. Then, click **OK**.

Step 3 Check whether the heap memory used by Flume reaches the threshold (95% of the maximum heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **Flume** > **Configuration**. On the page that is displayed, click **All Configurations** and choose **Flume** > **System**. Set **-Xmx** in the **GC_OPTS**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for the Flume server is insufficient for data transmission. You are advised to change the heap memory to: Channel capacity x Maximum size of a single data record x Number of channels. Note that the value of **xmx** cannot exceed the remaining memory of the node.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

- During the restart, the Flume service is interrupted.
- During the instance restart, if the failover mode of SinkGroup is configured and at least one instance is running properly, the Flume service is not interrupted. Otherwise, the Flume service is interrupted.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.215 ALM-24007 Flume Server Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the Flume service every 60 seconds. This alarm is generated when the direct memory usage of the Flume instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the Flume direct memory usage is less than or equal to the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24007	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the direct memory overflows, the service may break down and the Flume instance may be unavailable.

Possible Causes

The direct memory of the Flume process is overused or the direct memory is inappropriately allocated.

Handling Procedure

Check the direct memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **Flume Direct Memory Usage Exceeds the Threshold**, and view the **Location** information. Check the name of the host for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Flume**. On the page that is displayed, click the **Instance** tab. On the displayed tab page, select the role corresponding to the host name for which the alarm is generated and select **Customize** from the drop-down list in the upper right corner of the chart area. Choose **Agent** and select **Flume Direct Memory Resource Percentage**. Then, click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the direct memory used by Flume reaches the threshold (80% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **Flume** > **Configuration**. On the page that is displayed, click **All Configurations** and choose **Flume** > **System**. Set **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the direct memory size configured for the Flume server instance cannot meet service requirements. You are advised to change the value of **-XX:MaxDirectMemorySize** to twice the current direct memory size or change the value based on site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

- During the restart, the Flume service is interrupted.
- During the instance restart, if the failover mode of SinkGroup is configured and at least one instance is running properly, the Flume service is not interrupted. Otherwise, the Flume service is interrupted.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.216 ALM-24008 Flume Server Non-Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the non-heap memory usage of the Flume service every 60 seconds. This alarm is generated when the non-heap memory usage of the Flume instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the non-heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24008	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the non-heap memory overflows, the service may break down and the Flume instance may be unavailable.

Possible Causes

The non-heap memory of the Flume instance is overused or the non-heap memory is inappropriately allocated.

Handling Procedure

Check non-heap memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **Flume Non-Heap Memory Usage Exceeds the Threshold**, and view the **Location** information. Check the name of the host for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Flume**. On the page that is displayed, click the **Instance** tab. On the displayed tab page, select the role corresponding to the host name for which the alarm is generated and select **Customize** from the drop-down list in the upper right corner of the chart area. Choose **Agent** and select **Flume Non-Heap Memory Resource Percentage**. Then, click **OK**.

Step 3 Check whether the non-heap memory used by Flume reaches the threshold (80% of the maximum non-heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **Flume** > **Configuration**. On the page that is displayed, click **All Configurations** and choose **Flume** > **System**. Set **-XX: MaxPermSize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the non-heap memory size configured for the Flume server instance cannot meet service requirements. You are advised to change the value of **-XX:MaxPermSize** to twice the current non-heap memory size or change the value based on site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

- During the restart, the Flume service is interrupted.
- During the instance restart, if the failover mode of SinkGroup is configured and at least one instance is running properly, the Flume service is not interrupted. Otherwise, the Flume service is interrupted.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.217 ALM-24009 Flume Server Garbage Collection (GC) Time Exceeds the Threshold

Alarm Description

The system checks the GC duration of the Flume process every 60 seconds. This alarm is generated when the GC duration of the Flume process exceeds the threshold (12 seconds by default) for five consecutive times. This alarm is cleared when the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24009	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Flume data transmission efficiency decreases.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The heap memory of the Flume process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC duration.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **GC Duration Exceeds the Threshold**, and view the **Location** information. Check the name of the host for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the target cluster* > **Services** > **Flume**. On the page that is displayed, click the **Instance** tab. On the displayed tab page, select the role corresponding to the host name for which the alarm is generated and select **Customize** from the drop-down list in the upper right corner of the chart area. Choose **Agent** and select **Garbage Collection (GC) Duration of Flume**. Then, click **OK**.

Step 3 Check whether the GC duration of the Flume process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **Flume** > **Configuration**. On the page that is displayed, click **All Configurations** and choose **Flume** > **System**. Set **-Xmx** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for the Flume server is insufficient for data transmission. You are advised to change the heap memory to: Channel capacity x Maximum size of a single data record x Number of channels. Note that the value of **xmx** cannot exceed the remaining memory of the node.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

- During the restart, the Flume service is interrupted.
- During the instance restart, if the failover mode of SinkGroup is configured and at least one instance is running properly, the Flume service is not interrupted. Otherwise, the Flume service is interrupted.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.218 ALM-24010 Flume Certificate File Is Invalid or Damaged

This section applies to MRS 3.2.0 or later.

Alarm Description

Flume checks whether the Flume certificate file is valid (whether the certificate exists and whether the certificate format is correct) every hour. This alarm is generated when the certificate file is invalid or damaged. This alarm is automatically cleared when the certificate file becomes valid again.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24010	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The Flume client cannot access the Flume server.

Possible Causes

The Flume certificate file is invalid or damaged.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **ALM-24010 Flume Certificate File Is Invalid or Damaged**, and view the **Location** information. View the IP address of the instance for which the alarm is generated.

Check whether the certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **su - omm** command to switch to user **omm**.

Step 3 Run the following command to go to the Flume service certificate directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/conf

Step 4 Run the **ls -l** command to check whether the **flume_sChat.crt** file exists.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Run the **openssl x509 -in flume_sChat.crt -text -noout** command to check whether certificate details are displayed properly.

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 Run the following command to go to the Flume script directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/bin

Step 7 Run the following command to generate a new certificate file. Then check whether the alarm is automatically cleared one hour later.

sh geneJKS.sh -f *Custom certificate password of the Flume role on the server* **-g** *Custom certificate password of the Flume role on the client*

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The custom certificate passwords must meet the following complexity requirements:

- Contain at least four types of uppercase letters, lowercase letters, digits, and special characters.
- Contain 8 to 64 characters.
- Be changed periodically (for example, every three months), and certificates and trust lists are generated again to ensure security.

Step 8 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 9**.
- If no, no further action is required.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.219 ALM-24011 Flume Certificate File Is About to Expire

This section applies to MRS 3.2.0 or later.

Alarm Description

Flume checks whether the Flume certificate file is about to expire every hour. This alarm is generated when the remaining validity period is at most 30 days. This alarm is automatically cleared when the remaining validity period is greater than 30 days.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24011	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Currently, there is no impact on the system.

Possible Causes

The Flume certificate file is about to expire.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **ALM-24011 Flume Certificate Is About to Expire**, and view the **Location** information. View the IP address of the instance for which the alarm is generated.

Check whether the certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **su - omm** command to switch to user **omm**.

Step 3 Run the following command to go to the Flume service certificate directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/conf

Step 4 Run the following command to check the effective time and expiration time of the Flume user certificate:

openssl x509 -noout -text -in flume_sChat.crt

Step 5 Perform **Step 6** to **Step 7** during off-peak hours to update the certificate file as needed.

Step 6 Run the following command to go to the Flume script directory:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/bin

Step 7 Run the following command to generate a new certificate file. Then, check whether the alarm is automatically cleared one hour later.

sh geneJKS.sh -f *Custom certificate password of the Flume role on the server* **-g** *Custom certificate password of the Flume role on the client*

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

NOTE

The custom certificate passwords must meet the following complexity requirements:

- Contain at least four types of uppercase letters, lowercase letters, digits, and special characters.
- Contain 8 to 64 characters.
- Be changed periodically (for example, every three months), and certificates and trust lists are generated again to ensure security.

Step 8 Log in to the Flume node for which the alarm is generated as user **omm** and repeat **Step 6** to **Step 7**. Then, check whether the alarm is automatically cleared one hour later.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 10**.
- If no, no further action is required.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 12 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.220 ALM-24012 Flume Certificate File Has Expired

This section applies to MRS 3.2.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Description

Flume checks whether its certificate file in the system has expired every hour. This alarm is generated when the server certificate has expired. This alarm is automatically cleared when the certificate file becomes valid again.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24012	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The Flume client cannot access the Flume server.

Possible Causes

The Flume certificate file has expired.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **ALM-24012 Flume Certificate Has Expired**, and view the **Location** information. View the IP address of the instance for which the alarm is generated.

Check whether the certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **su - omm** command to switch to user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Run the following command to go to the Flume service certificate directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/conf

Step 4 Run the following command to check the effective time and expiration time of the HA user certificate to determine whether the certificate file is still in the validity period:

openssl x509 -noout -text -in flume_sChat.crt

- If yes, go to **Step 9**.
- If no, go to **Step 5**.

Step 5 Run the following command to go to the Flume script directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/bin

Step 6 Run the following command to generate a new certificate file. Then, check whether the alarm is automatically cleared one hour later.

sh geneJKS.sh -f *Custom certificate password of the Flume role on the server* **-g** *Custom certificate password of the Flume role on the client*

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

  NOTE

  The custom certificate passwords must meet the following complexity requirements:
  - Contain at least four types of uppercase letters, lowercase letters, digits, and special characters.
  - Contain 8 to 64 characters.
  - Be changed periodically (for example, every three months), and certificates and trust lists are generated again to ensure security.

Step 7 Log in to the Flume node for which the alarm is generated as user **omm** and repeat **Step 5** to **Step 6**. Then, check whether the alarm is automatically cleared one hour later.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 9**.
- If no, no further action is required.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Flume** for the target cluster.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.221 ALM-24013 Flume MonitorServer Certificate File Is Invalid or Damaged

This section applies to MRS 3.2.0 or later.

Alarm Description

MonitorServer checks whether its certificate file is valid (whether the certificate exists and whether the certificate format is correct) every hour. This alarm is generated when the certificate file is invalid or damaged. This alarm is automatically cleared when the certificate file becomes valid again.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24013	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The Flume client cannot access the Flume server.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The MonitorServer certificate file is invalid or damaged.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **ALM-24013 MonitorServer Certificate File Is Invalid or Damaged**, and view the **Location** information. View the IP address of the instance for which the alarm is generated.

Check whether the certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **su - omm** command to switch to user **omm**.

Step 3 Run the following command to go to the MonitorServer certificate file directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/conf

Step 4 Run the **ls -l** command to check whether the **ms_sChat.crt** file exists:

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Run the **openssl x509 -in ms_sChat.crt -text -noout** command to check whether certificate details are displayed.

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 Run the following command to go to the Flume script directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/bin

Step 7 Run the following command to generate a new certificate file. Then check whether the alarm is automatically cleared one hour later.

sh geneJKS.sh -m *Custom password of the MonitorServer certificate on the server* **-n** *Custom password of the MonitorServer certificate on the client*

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

NOTE

The custom certificate passwords must meet the following complexity requirements:

- Contain at least four types of uppercase letters, lowercase letters, digits, and special characters.
- Contain 8 to 64 characters.
- Be changed periodically (for example, every three months), and certificates and trust lists are generated again to ensure security.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 9**.
- If no, no further action is required.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Select **MonitorServer** in the required cluster for **Service**.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.222 ALM-24014 Flume MonitorServer Certificate Is About to Expire

This section applies to MRS 3.2.0 or later.

Alarm Description

MonitorServer checks whether its certificate file is about to expire every hour. This alarm is generated when the remaining validity period is at most 30 days. This alarm is automatically cleared when the remaining validity period is greater than 30 days.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24014	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Currently, there is no impact on the system.

Possible Causes

The MonitorServer certificate file is about to expire.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **ALM-24014 MonitorServer Certificate Is About to Expire**, and view the **Location** information. View the IP address of the instance for which the alarm is generated.

Check whether the certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **su - omm** command to switch to user **omm**.

Step 3 Run the following command to go to the MonitorServer certificate file directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/conf

Step 4 Run the following command to check the effective time and expiration time of the MonitorServer user certificate:

openssl x509 -noout -text -in ms_sChat.crt

Step 5 Perform **Step 6** to **Step 7** during off-peak hours to update the certificate file as needed.

Step 6 Run the following command to go to the Flume script directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/bin

Step 7 Run the following command to generate a new certificate file. Then, check whether the alarm is automatically cleared one hour later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

sh geneJKS.sh -m *Custom password of the MonitorServer certificate on the server* **-n** *Custom password of the MonitorServer certificate on the client*

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

NOTE

The custom certificate passwords must meet the following complexity requirements:

- Contain at least four types of uppercase letters, lowercase letters, digits, and special characters.
- Contain 8 to 64 characters.
- Be changed periodically (for example, every three months), and certificates and trust lists are generated again to ensure security.

Step 8 Log in to the Flume node for which the alarm is generated as user **omm** and repeat **Step 6** to **Step 7**. Then, check whether the alarm is automatically cleared one hour later.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 10**.
- If no, no further action is required.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Select **MonitorServer** in the required cluster for **Service**.

Step 12 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.223 ALM-24015 Flume MonitorServer Certificate File Has Expired

This section applies to MRS 3.2.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Description

MonitorServer checks whether its certificate file in the system has expired every hour. This alarm is generated when the server certificate has expired. This alarm is automatically cleared when the MonitorServer certificate file becomes valid again.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
24015	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The Flume client cannot access the Flume server.

Possible Causes

The MonitorServer certificate file has expired.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing **ALM-24015 MonitorServer Certificate Has Expired**, and view the **Location** information. View the IP address of the instance for which the alarm is generated.

Check whether the certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **su - omm** command to switch to user **omm**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Run the following command to go to the MonitorServer certificate file directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/conf

Step 4 Run the following command to check the effective time and expiration time of the user certificate to determine whether the certificate file is still in the validity period:

openssl x509 -noout -text -in ms_sChat.crt

- If yes, go to **Step 9**.
- If no, go to **Step 5**.

Step 5 Run the following command to go to the Flume script directory:

cd ${BIGDATA_HOME}/FusionInsight_Porter_*/install/FusionInsight-Flume-*/flume/bin

Step 6 Run the following command to generate a new certificate file. Then, check whether the alarm is automatically cleared one hour later.

sh geneJKS.sh -m *Custom password of the MonitorServer certificate on the server* **-n** *Custom password of the MonitorServer certificate on the client*

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

  NOTE

  The custom certificate passwords must meet the following complexity requirements:
  - Contain at least four types of uppercase letters, lowercase letters, digits, and special characters.
  - Contain 8 to 64 characters.
  - Be changed periodically (for example, every three months), and certificates and trust lists are generated again to ensure security.

Step 7 Log in to the Flume node for which the alarm is generated as user **omm** and repeat **Step 5** to **Step 6**. Then, check whether the alarm is automatically cleared one hour later.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 9**.
- If no, no further action is required.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Select **MonitorServer** in the required cluster for **Service**.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.224 ALM-25000 LdapServer Service Unavailable

Description

The system checks the LdapServer service status every 30 seconds. This alarm is generated when the system detects that both the active and standby LdapServer services are abnormal.

This alarm is cleared when the system detects that one or two LdapServer services are normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25000	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The running status of the component that depends on the LdapServer becomes faulty. As a result, Kerberos authentication fails in the cluster or OS user cache synchronization is abnormal, and component services are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The node where the LdapServer service locates is faulty.
- The LdapServer process is abnormal.

Procedure

Check whether the nodes where the two SlapdServer instances of the LdapServer service are located are faulty.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **LdapServer** > **Instance**. Go to the LdapServer instance page to obtain the host name of the node where the two SlapdServer instances reside.

Step 2 Choose **O&M > Alarm > Alarms**. On the **Alarm** page of the FusionInsight Manager system, check whether any alarm of **Node Fault** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Check whether the host name in the alarm is consistent with the **Step 1** host name.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 Handle the alarm according to "ALM-12006 Node Fault".

Step 5 Check whether **LdapServer Service Unavailable** is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether the LdapServer process is normal.

Step 6 Choose **O&M > Alarm > Alarms**. On the **Alarm** page of the FusionInsight Manager system, check whether any alarm of **Process Fault** exists.

- If yes, go to **Step 7**.
- If no, go to **Step 10**.

Step 7 Check whether the service and host name in the alarm are consistent with the LdapServer service and host name.

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

Step 8 Handle the alarm according to "ALM-12007 Process Fault".

Step 9 Check whether **LdapServer Service Unavailable** is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On the FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 11 Select **LdapServer** in the required cluster from the **Service**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.225 ALM-25004 Abnormal LdapServer Data Synchronization

Description

The system checks the LdapServer data every 30 seconds. This alarm is generated when the data on the active and standby LdapServers of Manager is inconsistent for 12 consecutive times. This alarm is cleared when the data on the active and standby LdapServers is consistent.

The system checks the LdapServer data every 30 seconds. This alarm is generated when the LdapServer data in the cluster is inconsistent with that on Manager for 12 consecutive times. This alarm is cleared when the data is consistent.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25004	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

LdapServer data inconsistency occurs because the LdapServer data in Manager is damaged or the LdapServer data in the cluster is damaged. The LdapServer process with damaged data cannot provide services externally, and the authentication functions of Manager and the cluster are affected.

Possible Causes

- The network of the node where the LdapServer process locates is faulty.
- The LdapServer process is abnormal.
- The OS restart damages data on LdapServer.
- The amount of Oldap data exceeds the threshold (10 MB by default).

Procedure

Check whether the network where the LdapServer nodes reside is faulty.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. Record the IP address of HostName in the alarm locating information as IP1 (if multiple alarms exist, record the IP addresses as IP1, IP2, and IP3 respectively).

Step 2 Contact O&M personnel and log in to the nodes corresponding to IP 1. Run the ping command to check whether the IP address of the management plane of the active OMS node can be pinged.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 Contact the network administrator to recover the network and check whether **Abnormal LdapServer Data Synchronization** is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the LdapServer processes are normal.

Step 4 On the **Alarm** page of FusionInsight Manager, check whether the **OLdap Resource Abnormal** exists.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Clear the alarm by following the steps provided in **ALM-12004 OLdap Resource Abnormal**.

Step 6 Check whether **Abnormal LdapServer Data Synchronization** is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 7**.

Step 7 On the **Alarm** page of FusionInsight Manager, check whether **Process Faul**t is generated for the LdapServer service.

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Handle the alarm according to **ALM-12007 Process Fault**.

Step 9 Check whether **Abnormal LdapServer Data Synchronization** is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether the LdapServer processes are normal.

Step 10 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Record the IP address of HostName in the alarm locating information as "IP1" (if multiple alarms exist, record the IP addresses as "IP1", "IP2", and "IP3" respectively). Choose **Cluster** > *Name of the desired cluster* > **Services** > **LdapServer** > **Configurations**. Record the port number of LdapServer as "PORT". (If the IP address in the alarm locating information is the IP address of the standby management node, choose **System** > **OMS** > **oldap** > **Modify Configuration** and record the listening port number of LdapServer.)

Step 11 Log in to the nodes corresponding to IP1 as user **omm**.

Step 12 Run the following command to check whether errors are displayed in the queried information.

ldapsearch -H ldaps://*IP1*:*PORT* -LLL -x -D cn=root,dc=hadoop,dc=com -W -b ou=Peoples,dc=hadoop,dc=com

After running the command, enter the **LDAP** administrator password. Contact the system administrator to obtain the password.

- If yes, go to **Step 13**.
- If no, go to **Step 15**.

Step 13 Recover the LdapServer and OMS nodes using data backed up before the alarm is generated.

NOTE

Use the OMS data and LdapServer data backed up at the same point in time to recover the data. Otherwise, the service and operation may fail. To recover data when services run properly, you are advised to manually back up the latest management data and then recover the data. Otherwise, Manager data produced between the backup point in time and the recovery point in time will be lost.

Step 14 Check whether alarm **Abnormal LdapServer Data Synchronization** is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Check whether the data volume of the Oldap exceeds the threshold (10 MB by default). (This step applies only to versions earlier than MRS 3.3.0. For MRS 3.3.0 and later versions, go to **Step 18**.)

Step 15 Log in to the active OMS node as user **omm**.

Step 16 Run the following command to check whether the directory contains **.mdb** files.

ll /srv/BigData/ldapData/oldap/data/

- If yes, check and record the size of the **.mdb** file and go to **Step 17**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 18**.

Step 17 Run the following command to view the Oldap configuration and record the value of **Map size** (the default value is **10485760** bytes, that is, 10 MB)

mdb_stat -e /srv/BigData/ldapData/oldap/data/

Check whether the size of the **.mdb** file with **Step 16** records reaches the value of **Map size.**

- If yes, contact O&M personnel.
- If no, go to **Step 18**.

Collect fault information.

Step 18 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 19 Select **LdapServer** in the required cluster and **OmsLdapServer** from the **Service**.

Step 20 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 21 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.226 ALM-25005 nscd Service Exception

Alarm Description

The system checks the status of the nscd service every 60 seconds. This alarm is generated when the nscd process fails to be queried for four consecutive times (three minutes) or users in LdapServer cannot be obtained.

This alarm is cleared when the process is restored and users in LdapServer can be obtained.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
25005	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The alarmed node may not be able to synchronize data from LdapServer. The **id** command may fail to obtain the LDAP data, affecting upper-layer services.

Possible Causes

- The nscd service is not started.
- The network is faulty, and cannot access the LDAP server.
- NameService is abnormal.
- Users cannot be queried because the OS executes commands too slowly.

Handling Procedure

Check whether the nscd service is started.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Record the IP address of **HostName** in **Location** of the alarm as **IP1** (if multiple alarms exist, record the IP addresses as **IP1**, **IP2**, and **IP3** respectively).

Step 2 Contact the O&M personnel to access the node using IP1 as user **root**. Run the **ps -ef | grep nscd** command on the node and check whether the **/usr/sbin/nscd** process is started.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 Run the **service nscd restart** command as user **root** to restart the nscd service. Then run the **ps -ef | grep nscd** command to check whether the nscd service is started.

- If yes, go to **Step 4**.
- If no, go to **Step 15**.

Step 4 Wait for 5 minutes and run the ps -ef | grep nscd command again as user root. Check whether the service exists.

- If yes, go to **Step 11**.
- If no, go to **Step 15**.

Check whether the network is faulty, and whether the LDAP server can be accessed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Log in to the alarmed node as user **root** and run the **ping** command to check whether the network connectivity between this node and the LdapServer node is normal.

- If yes, go to **Step 6**.
- If no, contact network administrators to troubleshoot the fault.

Check whether the NameService is normal.

Step 6 Log in to the alarmed node as user **root**. Run the **cat /etc/nsswitch.conf** command to check whether the **passwd**, **group**, **services**, **netgroup**, and **aliases** of NameService are correctly configured.

The correct parameter configurations are as follows:

passwd: **compat ldap**; **group**: **compat ldap**; **services**: **files ldap**; **netgroup**: **files ldap**; **aliases**: **files ldap**

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Log in to the alarmed node as user **root**. Run the **cat /etc/nscd.conf** command to check whether the **enable-cache passwd**, **positive-time-to-live passwd**, **enable-cache group**, and **positive-time-to-live group** in the configuration file are correctly configured.

The correct parameter configurations are as follows:

enable-cache passwd: **yes**; **positive-time-to-live passwd**: **600**; **enable-cache group**: **yes**; **positive-time-to-live group**: **3600**

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

Step 8 Run the **/usr/sbin/nscd -i group** and **/usr/sbin/nscd -i passwd** commands as user **root**. Wait for 2 minutes and run the **id admin** and **id backup/manager** commands to check whether results can be queried.

- If yes, go to **Step 11**.
- If no, go to **Step 15**.

Step 9 Run the **vi /etc/nsswitch.conf** command as user **root**. Correct the configurations in **Step 6** and save the file. Run the **service nscd restart** command to restart the nscd service. Wait for 2 minutes and run the **id admin** and **id backup/manager** commands to check whether results can be queried.

- If yes, go to **Step 11**.
- If no, go to **Step 15**.

Step 10 Run the **vi /etc/nscd.conf** command as user **root**. Correct the configurations in **Step 7** and save the file. Run the **service nscd restart** command to restart the nscd service. Wait for 2 minutes and run the **id admin** and **id backup/manager** commands to check whether results can be queried.

- If yes, go to **Step 11**.
- If no, go to **Step 15**.

Step 11 Log in to the FusionInsight Manager portal. Wait for 5 minutes and check whether the **nscd Service Exception** alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check whether frame freezing occurs when running a command in the operating system.

Step 12 Log in to the faulty node as user **root**, run the **id admin** command, and check whether the command execution takes a long time. If the command execution takes more than 3 seconds, the command execution is deemed to be slow.

- If yes, go to **Step 13**.
- If no, go to **Step 15**.

Step 13 Run the **cat /var/log/messages** command to check whether the nscd frequently restarts or the error information "Can't contact LDAP server" exists.

nscd exception example:

```
Feb 11 11:44:42 10-120-205-33 nscd: nss_ldap: failed to bind to LDAP server ldaps://10.120.205.55:21780: Can't contact LDAP server
Feb 11 11:44:43 10-120-205-33 ntpq: nss_ldap: failed to bind to LDAP server ldaps://10.120.205.55:21780: Can't contact LDAP server
Feb 11 11:44:44 10-120-205-33 ntpq: nss_ldap: failed to bind to LDAP server ldaps://10.120.205.92:21780: Can't contact LDAP server
```

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Run the **vi$BIGDATA_HOME/tmp/random_ldap_ip_order** command to modify the number at the end. If the original number is an odd number, change it to an even number. If the number is an even number, change it to an odd number.

Run the **vi /etc/ldap.conf** command to enter the editing mode, press **Insert** to start editing, and then change the first two IP addresses of the URI configuration item.

After the modification is complete, press **Esc** to exit the editing mode and enter **:wq!** to save the settings and exit.

Run the **service nscd restart** command to restart the nscd service. Wait 5 minutes and run the **id admin** command again. Check whether the command execution is slow.

- If yes, go to **Step 15**.
- If no, log in to other faulty nodes and repeat **Step 12** to **Step 14** to check whether the first LdapServer node in the URI before modifying **/etc/ldap.conf** is faulty. For example, check whether the service IP address is unreachable, the network delay is too long, or other abnormal software is deployed.

Collect the fault information.

Step 15 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 16 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **LdapClient** for the target cluster.

Step 17 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 18 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.227 ALM-25006 Sssd Service Exception

Description

The system checks the status of the sssd service every 60 seconds. This alarm is generated when the sssd process fails to be queried for four consecutive times (three minutes) or users in LdapServer cannot be obtained.

This alarm is cleared when the process is restored and users in LdapServer can be obtained.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25006	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.

Impact on the System

The alarmed node may not be able to synchronize data from LdapServer. The id command may fail to obtain the LDAP data, affecting upper-layer services.

Possible Causes

- The sssd service is not started or is incorrectly started.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The network is faulty and cannot access the LDAP server.
- NameService is abnormal.
- Users cannot be queried because the OS executes commands too slowly.

Procedure

Check whether the sssd service is correctly started.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms**. Find the IP address of **HostName** in **Location** of the alarm and record it as IP1 (if multiple alarms exist, record the IP addresses as IP1, IP2, and IP3 respectively).

Step 2 Contact the O&M personnel to access the node using IP1 as user root. Run the **ps -ef | grep sssd** command and check whether the **/usr/sbin/sssd** process is started.

- If the process is started, go to **Step 3**.
- If the process is not started, go to **Step 4**.

Step 3 Check whether the sssd process queried in **Step 2** has three subprocesses.

- If yes, go to **Step 5**.
- If no, go to **Step 4**.

Step 4 Run the **service sssd restart** command as user **root** to restart the sssd service. Then run the **ps -ef | grep sssd** command to check whether the sssd process is normal.

In the normal state, the **/usr/sbin/sssd** process has three subprocesses: **/usr/libexec/sssd/sssd_be**, **/usr/libexec/sssd/sssd_nss**, and **/usr/libexec/sssd/sssd_pam**.

- If it exists, go to **Step 9**.
- If it does not exist, go to **Step 13**.

Check whether the LDAP server can be accessed.

Step 5 Log in to the alarmed node as user **root**. Run the **ping** command to check the network connectivity between this node and the LdapServer node.

- If the network is normal, go to **Step 6**.
- If the network is faulty, contact network administrators to troubleshoot the fault.

Check whether NameService is normal.

Step 6 Log in to the alarmed node as user **root**. Run the **cat /etc/nsswitch.conf** command and check the **passwd** and **group** configurations of NameService.

The correct parameter configurations are as follows: **passwd: files sss** and **group: files sss**.

- If the configurations are correct, go to **Step 7**.
- If the configurations are incorrect, go to **Step 8**.

Step 7 Run the **/usr/sbin/sss_cache -G** and **/usr/sbin/sss_cache -U** commands as user **root**. Wait for 2 minutes and run the **id admin** and **id backup/manager** commands to check whether results can be queried.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If results are queried, go to **Step 9**.
- If no result is queried, go to **Step 13**.

Step 8 Run the **vi /etc/nsswitch.conf** command as user **root**. Correct the configurations in **Step 6** and save the file. Run the **service sssd restart** command to restart the sssd service. Wait for 2 minutes and run the **id admin** and **id backup/manager** commands to check whether results can be queried.

- If it exists, go to **Step 9**.
- If no result is queried, go to **Step 13**.

Step 9 Log in to the FusionInsight Manager portal. Wait for 5 minutes and check whether the **sssd Service Exception** alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm persists, go to **Step 10**.

Check whether frame freezing occurs when running a command in the operating system.

Step 10 Log in to the faulty node as user **root**, run the **id admin** command, and check whether the command execution takes a long time. If the command execution takes more than 3 seconds, the command execution is deemed to be slow.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Run the **cat /var/log/messages** command to check whether the sssd frequently restarts or the error information **Can't contact LDAP server** exists.

sssd restart example:

```
Feb  7 11:38:16 10-132-190-105 sssd[pam]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd[nss]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd[nss]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd[be[default]]: Shutting down
Feb  7 11:38:16 10-132-190-105 sssd: Starting up
Feb  7 11:38:16 10-132-190-105 sssd[be[default]]: Starting up
Feb  7 11:38:16 10-132-190-105 sssd[nss]: Starting up
Feb  7 11:38:16 10-132-190-105 sssd[pam]: Starting up
```

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 Run the **vi $BIGDATA_HOME/tmp/random_ldap_ip_order** command to modify the number at the end. If the original number is an odd number, change it to an even number. If the number is an even number, change it to an odd number.

Run the **vi /etc/sssd/sssd.conf** command to reverse the first two IP addresses of the **ldap_uri** configuration item, save the settings, and exit.

Run the **ps -ef | grep sssd** command to query the ID of the sssd process, kill it, and run the **/usr/sbin/sssd -D -f** command to restart the sssd service. Wait 5 minutes and run the **id admin** command again.

Check whether the command execution is slow.

- If yes, go to **Step 13**.
- If no, log in to other faulty nodes and run **Step 10** to **Step 12**. Collect logs and check whether the first ldapserver node in the ldap_uri before

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

modifying **/etc/sssd/sssd.conf** is faulty. For example, check whether the service IP address is unreachable, the network latency is too long, or other abnormal software is deployed.

Collect fault information.

Step 13 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 14 Select **LdapClient** in the required cluster from the **Service**.

Step 15 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.228 ALM-25007 Number of SlapdServer Connections Exceeds the Threshold

Alarm Description

The system checks the number of process connections on the SlapdServer node every 30 seconds and compares the actual number with the threshold. This alarm is generated when the number of process connections exceeds the threshold (**1000** by default) for multiple times (**5** by default).

Its **Trigger Count** is configurable. If **Trigger Count** is set to **1**, this alarm is cleared when the number of process connections is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the number of process connections is less than or equal to 90% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
25007	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

SlapdServer may respond slowly or become unavailable. Kerberos authentication times out or OS user caches cannot be synchronized. Component services are faulty.

Possible Causes

- There are too many SlapdServer connections.
- The alarm threshold or alarm trigger count is improperly configured.

Handling Procedure

Check whether there are too many SlapdServer process connections.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **LdapServer**.

Step 2 On the LdapServer dashboard page, observe the SlapdServer process connections and decrease the connections based on service requirements.

Step 3 Wait about 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 4 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, choose **LdapServer** > **Other** > **SlapdServer Service Connections**, and check whether the alarm trigger count and alarm threshold are set properly.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Change the trigger count and alarm threshold based on the actual number of process connections, and apply the changes.

Step 6 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **LdapServer** for the target cluster.

Step 9 Specify **Hosts** for collecting logs, which is optional. By default, all hosts are selected.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.229 ALM-25008 SlapdServer CPU Usage Exceeds the Threshold

Alarm Description

The system checks the CPU usage of the SlapdServer node every 30 seconds and compares the actual usage with the threshold. This alarm is generated when the SlapdServer CPU usage exceeds the threshold for multiple times (**5** by default).

Its **Trigger Count** is configurable. If **Trigger Count** is set to **1**, this alarm is cleared when the SlapdServer CPU usage is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the SlapdServer CPU usage is less than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
25008	Critical (default threshold: 85%) Major (default threshold: 75%)	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

SlapdServer may respond slowly or become unavailable. Kerberos authentication times out or OS user caches cannot be synchronized. Component services are faulty.

Possible Causes

- The alarm threshold or alarm trigger count is improperly configured.
- The CPU configuration cannot meet service requirements, and the CPU usage reaches the upper limit.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, choose **LdapServer** > **Other** > **SlapdServer Service Total CPU Percentage**, and check whether the alarm trigger count and alarm threshold are set properly.

- If yes, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 2**.

Step 2 Change the trigger count and alarm threshold based on the actual CPU usage, and apply the changes.

Step 3 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the CPU usage reaches the upper limit.

Step 4 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the right pane, click this alarm and obtain the host name in **Location**.

Step 5 Choose **Cluster** > **Services** > **LdapServer**, click the **Instance** tab, and click the SlapdServer instance corresponding to the host name in **Step 4**.

Step 6 On the dashboard of the instance, observe the real-time data of the **CPU Usage of a Single SlapdServer Instance** chart for about 5 minutes and check whether the CPU usage exceeds the threshold (**75%** by default) for multiple times.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Check whether the status of other SlapdServer instances is normal. For details, see **Step 5** to **Step 6**.

- If yes, contact the MRS cluster administrator to evaluate whether to expand the capacity of SlapdServer instances. Then, go to **Step 8**.
- If no, repair the faulty SlapdServer instance and go to **Step 8**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **LdapServer** for the target cluster.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.230 ALM-25500 KrbServer Service Unavailable

Description

The system checks the KrbServer service status every 30 seconds. This alarm is generated when the system detects that the KrbServer service is abnormal.

This alarm is cleared when the system detects that the KrbServer service is normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25500	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The running status of the component that depends on the KrbServer becomes faulty. As a result, the Kerberos authentication of the cluster fails, and the component services are abnormal.

Possible Causes

- The node where the KrbServer service locates is faulty.
- The OLdap service is abnormal.

Procedure

Check whether the node where the KrbServer service locates is faulty.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **KrbServer** > **Instance**. Go to the KrbServer instance page and view the host name of the node where the KrbServer service is deployed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On the **Alarm** page of the FusionInsight Manager system, check whether any alarm of **Node Fault** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Check whether the host name in the alarm is consistent with the **Step 1** host name.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 Handle the alarm according to "ALM-12006 Node Fault".

Step 5 Check whether **KrbServer Service Unavailable** is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the OLdap service is normal.

Step 6 On the **Alarm** page of the FusionInsight Manager system, check whether any alarm of **OLdap Resource Abnormal** exists.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Handle the alarm according to "ALM-12004 OLdap Resource Abnormal".

Step 8 Check whether **KrbServer Service Unavailable** is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On the FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 10 Select **KrbServer** in the required cluster from the **Service**.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.231 ALM-25501 Too Many KerberosServer Requests

Alarm Description

The system checks the number of requests processed by the KerberosServer node every 30 seconds. This alarm is generated when the number of requests exceeds the threshold for multiple consecutive times (5 by default).

Its **Trigger Count** is configurable. If **Trigger Count** is set to **1**, this alarm is cleared when the number of process connections is less than or equal to the threshold. If **Trigger Count** is greater than **1**, this alarm is cleared when the number of requests is less than or equal to 90% of the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
25501	Critical (default threshold: 15,000) Major (default threshold: 10,000)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Details	Specifies alarm details.

Impact on the System

KerberosServer responds slowly. As a result, Kerberos authentication times out and component services are in error.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- There are too many KerberosServer requests.
- The alarm threshold or alarm trigger count is improperly configured.

Handling Procedure

Check whether there are too many KerberosServer requests.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **KrbServer** to go to the KrbServer overview page.

Step 2 Observe the **Total KerberosServer Requests** chart and reduce the number of KerberosServer authentication requests based on site requirements. If the number of KerberosServer authentication requests cannot be determined, contact the O&M personnel.

If no chart is available, click the drop-down arrow on the right, select **Customize**, select the desired item, and click **OK**.

Step 3 Wait about 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 4 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, choose **KrbServer** > **Other** > **Total KerberosServer Requests**, and check whether the alarm trigger count and alarm threshold are set properly.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

Step 5 Change the trigger count and alarm threshold based on the actual number of requests, and apply the changes.

Step 6 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **KrbServer** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.232 ALM-26051 Storm Service Unavailable

Description

The system checks the Storm service status every 30 seconds. This alarm is generated when all Nimbus nodes in the cluster are abnormal and the Storm service is unavailable.

This alarm is cleared when the Storm service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
26051	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The cluster cannot provide the Storm service, and users cannot perform new Storm tasks.

Possible Causes

- The Kerberos cluster is faulty.
- The ZooKeeper cluster is faulty or suspended.
- The active and standby Nimbus nodes in the Storm cluster are abnormal

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the status of the Kerberos cluster. (Skip this step if the normal mode is used.)

Step 1 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services**.

Step 2 Check whether the running status of the Kerberos service is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 See the related maintenance information of **ALM-25500 KrbServer Service Unavailable**.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the status of the ZooKeeper cluster.

Step 5 Check whether the running status of the ZooKeeper service is **Normal**.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 If ZooKeeper service is stopped, start it, else see the related maintenance information of **ALM-13000 ZooKeeper Service Unavailable**.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check the status of the active and standby Nimbus nodes.

Step 8 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Nimbus** to go to the Nimbus Instances page.

Step 9 Check whether only one Nimbus node that is in the **Active** state in **Roles**.

- If yes, go to **Step 13**.
- If no, go to **Step 10**.

Step 10 Select two Nimbus role instances, choose **More** > **Restart Instance**, and check whether the instances restart successfully.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Log in to the FusionInsight Manager portal again, choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Nimbus** to check whether the running status is **Normal**.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 Wait for 30 seconds and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collecting Fault Information

Step 13 On the FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 14 Select the following nodes in the required cluster from the **Service** drop-down list:

- KrbServer

  NOTE

  KrbServer logs do not need to be downloaded in normal mode.
- ZooKeeper
- Storm

Step 15 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.233 ALM-26052 Number of Available Supervisors of the Storm Service Is Less Than the Threshold

Description

The system periodically checks the number of available Supervisors every 60 seconds and compares the number of available Supervisors with the threshold. This alarm is generated when the number of available Supervisors is less than the threshold.

You can change the threshold in **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster*.

This alarm is cleared when the number of available Supervisors is greater than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
26052	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Existing tasks in the cluster cannot be performed. The cluster can receive new Storm tasks, but cannot perform these tasks.

Possible Causes

The status of some Supervisors in the cluster is abnormal.

Procedure

Check the Supervisor status.

Step 1 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Supervisor** to go to the Storm service management page.

Step 2 In **Roles**, check whether any instance whose status is **Faulty** or **Restoring** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Select Supervisor role instances whose status is **Faulty** or **Restoring**, choose **More** > **Restart Instance**, and check whether the instances restart successfully.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Wait for 30 seconds, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

Services are interrupted when the Supervisor is being restarted. Then, services are restored after the restarting.

Collect fault information.

Step 5 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 6 Select **Storm** and **ZooKeeper** in the required cluster from the **Service** drop-down list box.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.234 ALM-26053 Storm Slot Usage Exceeds the Threshold

Description

The system checks the slot usage every 60 seconds and compares the actual slot usage with the threshold. This alarm is generated when the slot usage is greater than the threshold.

You can change the threshold in **O&M** > **Alarm** > **Thresholds**.

This alarm is cleared when the slot usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
26053	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

New Storm tasks cannot be performed.

Possible Causes

- The status of some Supervisors in the cluster is abnormal.
- The status of all Supervisors is normal, but the processing capability is insufficient.

Procedure

Check the Supervisor status.

Step 1 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Instance** to go to the Storm instance management page.

Step 2 Check whether any instance whose status is **Faulty** or **Restoring** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Select Supervisor role instances whose status is **Faulty** or **Restoring**, choose **More** > **Restart Instance**, and check whether the instances restart successfully.

- If yes, go to **Step 4**.
- If no, go to **Step 10**.

Step 4 Wait several minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Increase the number of slots in each Supervisor.

Step 5 Log in to the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Configurations** > **All Configurations**.

Step 6 Increase the number of ports in the **supervisor.slots.ports** parameter of each Supervisor role and restart the instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Wait several minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 8 Expand the capacity of the Supervisors.

Step 9 Wait several minutes, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

  NOTE

  Services are interrupted when the Supervisor is being restarted. Then, services are restored after the restarting.

Collect fault information.

Step 10 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 11 Select **Storm** and **ZooKeeper** in the required cluster from the **Service** drop-down list box.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.235 ALM-26054 Nimbus Heap Memory Usage Exceeds the Threshold

Description

The system checks the heap memory usage of Storm Nimbus every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the heap memory usage of Storm Nimbus exceeds the threshold (80% of the maximum memory by default) for 5 consecutive times.

Users can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Storm** > **Nimbus** to change the threshold.

The alarm is cleared when the heap memory usage is less than or equal to the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
26054	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

When the heap memory usage of Storm Nimbus is overhigh, frequent GCs occur. In addition, a memory overflow may occur so that the Yarn service is unavailable.

Possible Causes

The heap memory of the Storm Nimbus instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Check the heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Heap Memory Usage of Storm Nimbus Exceeds the Threshold** > **Location**. Check the host name of the instance for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down menu in the chart area and choose **Customize** > **Nimbus** > **Heap Memory Usage of Nimbus**. Click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the used heap memory of Nimbus reaches the threshold (The default value is 80% of the maximum heap memory) specified for Nimbus.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Storm** > **Configurations** > **All Configurations** > **Nimbus** > **System**. Change the value of **-Xmx** in **NIMBUS_GC_OPTS** based on site requirements, and click **Save**. Click **OK**.

NOTE

- You are advised to set **-Xms** and **-Xmx** to the same value to prevent adverse impact on performance when JVM dynamically adjusts the heap memory size.
- The number of Workers grows as the Storm cluster scale increases. You can increase the value of **GC_OPTS** for Nimbus. The recommended value is as follows: If the number of Workers is 20, set **-Xmx** to a value greater than or equal to 1 GB. If the number of Workers exceeds 100, set **-Xmx** to a value greater than or equal to 5 GB.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTE

During the restart of the service or instance, services are interrupted. After the service or instance is restarted, services are restored.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select the following node in the required cluster from the **Service** drop-down list.

- NodeAgent
- Storm

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.236 ALM-27001 DBService Service Unavailable

Description

The alarm module checks the DBService service status every 30 seconds. This alarm is generated when the system detects that DBService service is unavailable.

This alarm is cleared when DBService service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
27001	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The database service is unavailable and cannot provide data import and query functions for upper-layer services, which results in some services exceptions.

Possible Causes

- The floating IP address does not exist.
- There is no active DBServer instance.
- The active and standby DBServer processes are abnormal.

Procedure

Check whether the floating IP address exists in the cluster environment.

Step 1 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService** > **Instance**.

Step 2 Check whether the active instance exists.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 3**.
- If no, go to **Step 9**.

Step 3 Select the active DBServer instance and record the IP address.

Step 4 Log in to the host that corresponds to the preceding IP address as user **root**, and run the **ifconfig** command to check whether the DBService floating IP address exists on the node.

- If yes, go to **Step 5**.
- If no, go to **Step 9**.

Step 5 Run the **ping** *floatip* command to check whether the DBService floating IP address can be pinged successfully.

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 Log in to the host that corresponds to the DBService floating IP address as user **root**, and run the command to delete the floating IP address.

ifconfig *interface* **down**

Step 7 On the FusionInsight Manager home page, choose **Cluster > Services** > **DBService** > **More** > **Restart Service** to restart DBService, and check whether DBService is restarted successfully.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Wait for about 2 minutes and check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 14**.

Check the status of the active DBServer instance.

Step 9 Select the DBServer instance whose role status is abnormal and record the IP address.

Step 10 On the **Alarm** page, check whether **Process Fault** occurs in the DBServer instance on the host that corresponds to the IP address.

- If yes, go to **Step 11**.
- If no, go to **Step 19**.

Step 11 Handle the alarm according to "ALM-12007 Process Fault".

Step 12 Wait for about 5 minutes and check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 19**.

Check the status of the active and standby DBServers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Log in to the host that corresponds to the preceding IP address as user **root**, and run the **su - omm** command to switch to user **omm**.

Step 14 Run the **cd ${DBSERVER_HOME}** command to go to the installation directory of the DBService.

Step 15 Run the **sh sbin/status-dbserver.sh** command to view the status of the active and standby HA processes of DBService. Determine whether the status can be viewed successfully.

```
HAMode
double

NodeName            HostName        HAVersion          StartTime           HAActive
HAAllResOK          HARunPhase
10_5_89_12          host01          V100R001C01        2019-06-13 21:33:09 active
normal              Actived
10_5_89_66          host03          V100R001C01        2019-06-13 21:33:09 standby
normal              Deactived

NodeName            ResName         ResStatus          ResHAStatus         ResType
10_5_89_12          floatip         Normal             Normal              Single_active
10_5_89_12          gaussDB         Active_normal      Normal              Active_standby
10_5_89_66          floatip         Stopped            Normal              Single_active
10_5_89_66          gaussDB         Standby_normal     Normal              Active_standby
```

- If yes, go to **Step 16**.
- If no, go to **Step 19**.

Step 16 Check whether the active and standby HA processes are in the abnormal state.

- If yes, go to **Step 17**.
- If no, go to **Step 19**.

Step 17 On FusionInsight Manager, choose **Cluster > Services** > **DBService** > **More** > **Restart Service** to restart DBService, and check whether the system displays a message indicating that the restart is successful.

- If yes, go to **Step 18**.

  NOTE

  The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

- If no, go to **Step 19**.

Step 18 Wait for about 2 minutes and check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 19**.

Collect fault information.

Step 19 On FusionInsight Manager, choose **O&M** > **Log > Download**.

Step 20 Select **DBService** in the required cluster and **NodeAgent** from the **Service**.

Step 21 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 22 Contact the O&M personnel and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.237 ALM-27003 DBService Heartbeat Interruption Between the Active and Standby Nodes

Alarm Description

This alarm is generated when the active or standby DBService node does not receive heartbeat messages from the peer node for 7 seconds.

This alarm is cleared when the heartbeat recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
27003	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Local DBService HA Name	Specifies a local DBService HA.
Peer DBService HA Name	Specifies a peer DBService HA.

Impact on the System

During the DBService heartbeat interruption, only one node can provide the service. If this node is faulty, no standby node is available for failover and the service is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The link between the active and standby DBService nodes is abnormal.

Handling Procedure

Check whether the network between the active DBService server and the standby DBService server is normal.

Step 1 In the alarm list on FusionInsight Manager, click ⌄ in the row where the alarm is located in the real-time alarm list and view the address of the standby DBService node.

Step 2 Log in to the active DBService node as user **root**.

To obtain the IP address of the active DBService node, log in to Manager and choose **Cluster** > **Services** > **DBService** > **Instances**.

Step 3 Run the following command to check whether the standby DBService server is accessible:

```
ping heartbeat IP address of the standby DBService server
```

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 7 Select the following nodes in the required cluster from the **Service**:

- DBService
- Controller
- NodeAgent

Step 8 Click ✎ in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 9 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.238 ALM-27004 Data Inconsistency Between Active and Standby DBServices

Description

The system checks the data synchronization status between the active and standby DBService every 10 seconds. This alarm is generated when the synchronization status cannot be queried for six consecutive times or when the synchronization status is abnormal.

This alarm is cleared when the synchronization status becomes normal.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
27004	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Local DBService HA Name	Specifies the HA name of the local DBService.
Peer DBService HA Name	Specifies the HA name of the peer DBService.
SYNC_PERCENT	Specifies the synchronization percentage.

Impact on the System

When data is not synchronized between the active and standby DBServer and the active instance becomes abnormal, the data may be lost or abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The network between the active and standby nodes is unstable.
- The standby DBService is abnormal.
- The standby node disk space is full.
- The CPU usage of the GaussDB process on the active DBService node is high. You need to locate the failure cause based on logs.

Procedure

Check whether the network between the active and standby nodes is normal.

Step 1 On FusionInsight Manager, choose **Cluster > Services > DBService > Instance**, check the service IP address of the standby DBServer instance.

Step 2 Log in to the active DBService node as user **root**.

Step 3 Run the **ping** *Standby DBService heartbeat IP address* command to check whether the standby DBService node is reachable.

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the standby DBService is normal.

Step 6 Log in to the standby DBService node as user **root**.

Step 7 Run the **su - omm** command to switch to user **omm**.

Step 8 Go to the **${DBSERVER_HOME}/sbin** directory and run the **./status-dbserver.sh** command to check whether the GaussDB resource status of the standby DBService is normal. In the command output, check whether the following information is displayed in the row where **ResName** is **gaussDB**:

For example:

```
10_10_10_231 gaussDB Standby_normal Normal Active_standby
```

- If yes, go to **Step 9**.
- If no, go to **Step 16**.

Check whether the standby node disk space is full. (Skip this check for versions later than MRS 3.1.2.)

Step 9 Log in to the standby DBService node as user **root**.

Step 10 Run the **su - omm** command to switch to user **omm**.

Step 11 Go to the **${DBSERVER_HOME}** directory, and run the following commands to obtain the DBService data directory:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

cd ${DBSERVER_HOME}

source .dbservice_profile

echo ${DBSERVICE_DATA_DIR}

Step 12 Run the **df -h** command to view the system disk partition usage information.

Step 13 Check whether the DBService data directory space is full.

- If yes, go to **Step 14**.
- If no, go to **Step 16**.

Step 14 Expand the disk capacity of the node.

Step 15 After the disk capacity is expanded, wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Collect fault information.

Step 16 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 17 In the **Service** area, select **DBService** of the target cluster and **OS**, **OS Statistics**, and **OS Performance** under **OMS**, and click **OK**.

Step 18 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 19 Contact the O&M personnel and send the collected logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.239 ALM-27005 Database Connections Usage Exceeds the Threshold

Description

The system checks the usage of the number of database connections of the nodes where DBServer instances are located every 30 seconds and compares the usage with the threshold. If the usage exceeds the threshold for five consecutive times (this number is configurable, and 5 is the default value), the system generates this alarm. The default usage threshold is 90%, and you can configure it based on site requirements.

The trigger count is configurable. This alarm is cleared in the following scenarios:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The trigger count is 1, and the usage of the number of database connections is less than or equal to the threshold.
- The trigger count is greater than 1, and the usage of the number of database connections is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
27005	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Upper-layer services may fail to connect to the DBService database, affecting services.

Possible Causes

- Too many database connections are used.
- The maximum number of database connections is improperly configured.
- The alarm threshold or alarm trigger count is improperly configured.

Procedure

Checking whether too many data connections are used

Step 1 On FusionInsight Manager, click DBService in the service list on the left navigation pane. The DBService monitoring page is displayed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Observe the number of connections used by the database user. Based on the service scenario, reduce the number of database user connections.

Step 3 Wait for 2 minutes and check whether the alarm is automatically cleared.

- If it is, no further action is required.
- If it is not, go to **Step 4**.

Checking whether the maximum number of database connections is properly configured

Step 4 Log in to FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService** > **Configurations**. On the displayed page, select the **All Configurations** tab, and increase the maximum number of database connections based on service requirements. Click **Save**. In the displayed **Save configuration** dialog box, click **OK**.

Step 5 After the maximum number of database connections is changed, restart DBService (do not restart the upper-layer services).

Procedure: Log in to FusionInsight Manager and choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService**. On the displayed page, choose **More** > **Restart Service**. Enter the password of the current login user and click **OK**. Do not select **Restart upper-layer services.**, click **OK**.

NOTE

The service will be unavailable during the restart. In addition, upper-layer services that depend on the service are affected.

Step 6 After the service is restarted, wait for 2 minutes and check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 7**.

Checking whether the alarm threshold or trigger count is properly configured

Step 7 Log in to FusionInsight Manager and change the alarm threshold and alarm trigger count based on the actual database connection usage. Choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **DBService > Database > Database Connections Usage (DBServer)**. In the **Database Connections Usage (DBServer)** area, click the pencil icon next to **Trigger Count**. In the displayed dialog box, change the trigger count.

NOTE

Trigger Count: If the usage of the number of database connections exceeds the threshold consecutively for more than the value of this parameter, an alarm is generated.

Based on the actual database connection usage, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **DBService > Database > Database Connections Usage (DBServer)**. In the **Database Connections Usage (DBServer)** area, click **Modify** in the **Operation** column. In the **Modify Rule** dialog box, modify the required parameters and click **OK**.

Step 8 Wait for 2 minutes and check whether the alarm is automatically cleared.

- If it is, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If it is not, go to **Step 9**.

Collect fault information

Step 9 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 10 Select **DBService** in the required cluster from the **Service**.

Step 11 Specify the host for collecting logs by setting the **Host** parameter that is optional. By default, all hosts are selected.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.240 ALM-27006 Disk Space Usage of the Data Directory Exceeds the Threshold

Description

The system checks the disk space usage of the data directory on the active DBServer node every 30 seconds and compares the disk usage with the threshold. The alarm is generated when the disk space usage exceeds the threshold for five consecutive times (the default value). The number of consecutive times is configurable. The disk space usage threshold of the data directory is set to 80% by default, which is configurable as well.

The value of **hit number** is configurable. When the value is set to **1** and the disk space usage is lower than or equal to the threshold, the alarm is cleared. When the value is greater than 1 and the disk space usage is lower than 90% of the threshold, the alarm is cleared.

Attribute

Alarm ID	Alarm Severity	Auto Clear
27006	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
ClusterName	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the disk partition where the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the actual indicator value exceeds this threshold, the alarm is generated.

Impact on the System

- The DBService service process cannot provide the API for data writing.
- When the disk space usage of the data directory exceeds 90%, the database enters the read-only mode and "Database Enters the Read-Only Mode" is generated. As a result, service data cannot be written to the database.

Possible Causes

- The alarm threshold is improperly configured.
- The data volume of the database is too large or the disk configuration cannot meet service requirements, causing excessive disk usage.

Procedure

Check whether the threshold is set properly.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **DBService** > **Database** > **Disk Space Usage of the Data Directory** to check whether the alarm threshold is proper (the default value 80% is a proper value).

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Change the alarm threshold based on the actual service situation.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService**. On the **Dashboard** page, view the **Disk Space Usage of the Data Directory** chart and check whether the disk space usage of the data directory is lower than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether large files are incorrectly written into the disk.

Step 5 Log in to the active DBService node as user **omm**.

Step 6 Run the following commands to view the files whose size exceeds 500 MB in the data directory and check whether there are large files incorrectly written into the directory:

source $DBSERVER_HOME/.dbservice_profile

find "$DBSERVICE_DATA_DIR"/../ -type f -size +500M

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Handle the large files based on the actual scenario and check whether the alarm is cleared 2 minutes later.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **DBService** for the target cluster.

Step 10 Specify the host for collecting logs by setting the **Host** parameter which is optional. By default, all hosts are selected.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.241 ALM-27007 Database Enters the Read-Only Mode

Description

The system checks the disk space usage of the data directory on the active DBServer node every 30 seconds. The alarm is generated when the disk space usage exceeds 90%.

The alarm is cleared when the disk space usage is lower than 80%.

Attribute

Alarm ID	Alarm Severity	Auto Clear
27007	Critical	Yes

Parameters

Name	Meaning
ClusterName	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the actual indicator value exceeds this threshold, the alarm is generated.

Impact on the System

- Service data is lost.
- Data cannot be written for upper-layer services and the data is lost.

Possible Causes

The disk configuration cannot meet service requirements. The disk usage reaches the upper limit.

Procedure

Check whether the disk space usage reaches the upper limit.

Step 1 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On the **Dashboard** page, view the **Disk Space Usage of the Data Directory** chart and check whether the disk space usage of the data directory exceeds 90%.

- If yes, go to **Step 3**.
- If no, go to **Step 13**.

Step 3 Log in to the active management node of the DBServer as user **omm** and run the following commands to check whether the database enters the read-only mode:

source $DBSERVER_HOME/.dbservice_profile

gsql -U omm -W *password* **-d postgres -p 20051**

show default_transaction_read_only;

NOTE

In the preceding commands, *password* indicates the password of user **omm** of the DBService database. You can run the **\q** command to exit the database.

Check whether the value of **default_transaction_read_only** is **on**.

```
POSTGRES=# show default_transaction_read_only;
 default_transaction_read_only
-------------------------------
 on
(1 row)
```

- If yes, go to **Step 4**.
- If no, go to **Step 13**.

Step 4 Run the following commands to open the **dbservice.properties** file:

source $DBSERVER_HOME/.dbservice_profile

vi ${DBSERVICE_SOFTWARE_DIR}/tools/dbservice.properties

Step 5 Change the value of **gaussdb_readonly_auto** to **OFF**.

Step 6 Run the following command to open the **postgresql.conf** file:

vi ${DBSERVICE_DATA_DIR}/postgresql.conf

Step 7 Delete **default_transaction_read_only = on**.

Step 8 Run the following command for the configuration to take effect:

gs_ctl reload -D ${DBSERVICE_DATA_DIR}

Step 9 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. On the right of the alarm "Database Enters the Read-Only Mode", click **Clear** in the **Operation** column. In the dialog box that is displayed, click **OK** Manually clear the alarm.

Step 10 Log in to the active management node of the DBServer as user **omm** and run the following commands to view the files whose size exceeds 500 MB in the data directory and check whether there are large files incorrectly written into the directory:

source $DBSERVER_HOME/.dbservice_profile

find "$DBSERVICE_DATA_DIR"/../ -type f -size +500M

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Handle the files that are incorrectly written into the directory based on the actual scenario.

Step 12 Log in to FusionInsight Manager and choose **Cluster** > *Name of the desired cluster* > **Services** > **DBService**. On the **Dashboard** page, view the **Disk Space Usage of the Data Directory** chart and check whether the disk space usage is lower than 80%.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **DBService** for the target cluster.

Step 15 Specify the host for collecting logs by setting the **Host** parameter which is optional. By default, all hosts are selected.

Step 16 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 17 Contact O&M personnel and send the collected fault logs.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.242 ALM-29000 Impala Service Unavailable

Alarm Description

The alarm module checks the Impala service status every 30 seconds. This alarm is generated if the Impala service is abnormal.

This alarm is cleared after the Impala service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29000	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

When the Impala service is abnormal, you cannot perform cluster operations on Impala through FusionInsight Manager. The Impala service functions are unavailable.

Possible Causes

- The Hive service is abnormal.
- The KrbServer service is abnormal.
- The Impala process is abnormal.
- There are too many JDBC or ODBC connections.

Handling Procedure

Check whether the services on which Impala depends are normal.

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** to check whether Hive and KrbServer are stopped.

- If yes, start the stopped services and go to **Step 2**.
- If no, go to **Step 3**.

Step 2 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, check whether the **Impala Service Unavailable** alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, check whether ALM-16004 Hive Service Unavailable and ALM-25500 KrbServer Service Unavailable exist.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Rectify the fault by following the handling procedure of "ALM-16004 Hive Service Unavailable" and "ALM-25500 KrbServer Service Unavailable". Then, check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the Impala process is normal.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. Check whether ALM-12007 Process Fault exists in the alarm list.

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 Rectify the fault by following the handling procedure of "ALM-12007 Process Fault", and then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check whether there are too many JDBC or ODBC connections.

Step 7 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, check whether "ALM-29005 Number of JDBC Connections to Impalad Exceeds the Threshold" or "ALM-29006 Number of ODBC Connections to Impalad Exceeds the Threshold" exists.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Rectify the fault by referring to the handling procedure of "ALM-29005 Number of JDBC Connections to Impalad Exceeds the Threshold" or "ALM-29006 Number of ODBC Connections to Impalad Exceeds the Threshold". Then, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

After the fault that triggers the alarm is rectified, the alarm is automatically cleared.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.243 ALM-29004 Impalad Process Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the Impalad process every 30 seconds. This alarm is generated when the system detects that the memory usage exceeds the default threshold (80%).

This alarm is automatically cleared when the system detects that the memory usage of the process falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29004	Minor	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The memory usage is too high. Some query tasks may fail due to insufficient memory.

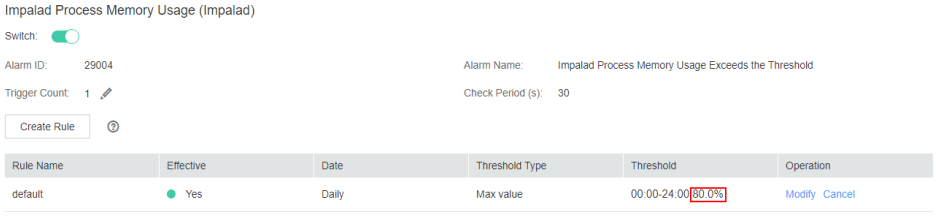
Possible Causes

The Impalad process is executing a large number of query tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **CPU and Memory** > **Impalad Process Memory Usage (Impalad)** and check the threshold.



Step 2 If the alarm threshold is smaller than 80%, increase the alarm threshold as required and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 If the threshold is greater than 80%, check whether a large number of concurrent query tasks exist when the alarm is generated. A large number of concurrent query tasks will cause the memory usage to increase sharply. After the tasks are complete, check whether the alarm is automatically cleared. During this period, some tasks may fail to be executed or may be canceled due to insufficient memory. In this case, try again.

NOTE

If the memory usage always exceeds the threshold, the cluster capacity needs to be expanded.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

The alarm is automatically cleared after the burst concurrent tasks are complete.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.244 ALM-29005 Number of JDBC Connections to Impalad Exceeds the Threshold

Alarm Description

The system checks the number of client connections to the Impalad node every 30 seconds. This alarm is generated when the number of client connections exceeds the customized threshold (60 by default).

This alarm is automatically cleared when the number of client connections is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29005	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

New client connections may be blocked or even fail.

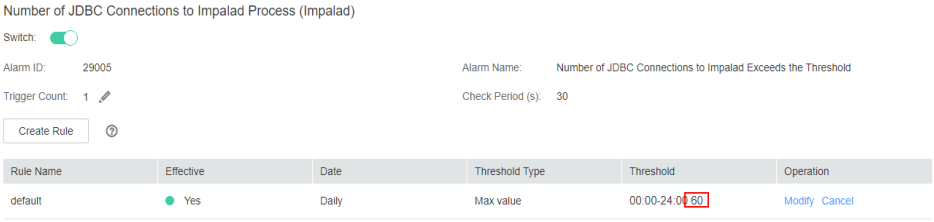
Possible Causes

Too many connections have been established with the Impala server or the threshold is too small.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

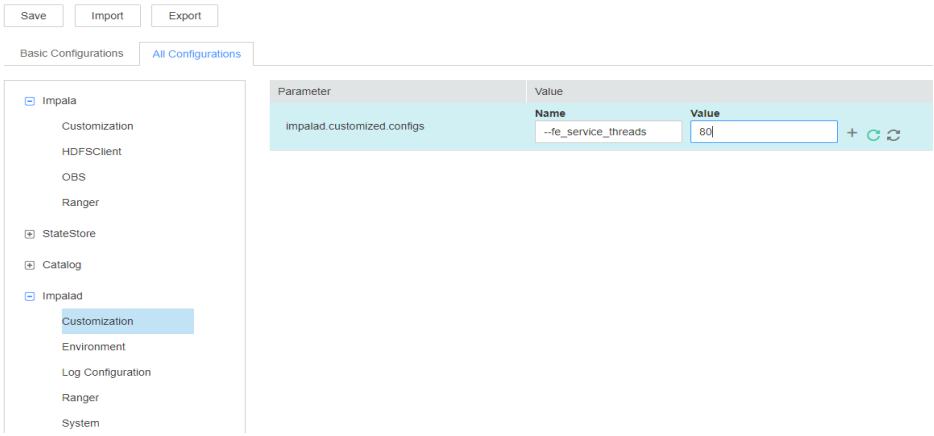
Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Connections** > **Number of JDBC Connections to Impalad Process** to check the configured threshold.



Step 2 Check the number of JDBC applications connected to Impalad and stop idle applications. Check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3** to change the number of concurrent client connections.

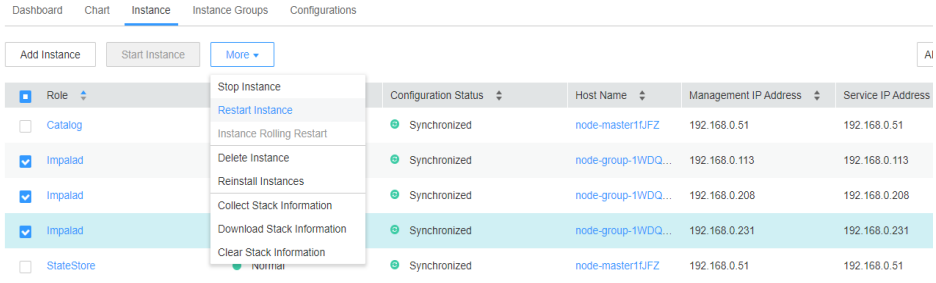
Step 3 On FusionInsight Manager, choose **Cluster** > **Impala** > **Configurations** > **All Configurations** > **Impalad** > **Customization**. Add the customized parameter **--fe_service_threads**. The default value of this parameter is **64**. Change the value as required and click **Save**.



Step 4 After the query tasks on all clients are complete, click the **Instances** tab. Select all Impalad instances, and restart them.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 After the restart is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

After the fault that triggers the alarm is rectified, the alarm is automatically cleared.

Related Information

None

9.11.245 ALM-29006 Number of ODBC Connections to Impalad Exceeds the Threshold

Alarm Description

The system checks the number of client connections to the Impalad node every 30 seconds. This alarm is generated when the number of client connections exceeds the customized threshold (60 by default).

This alarm is automatically cleared when the number of client connections is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29006	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

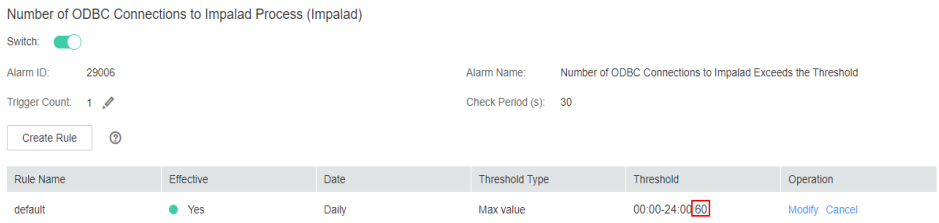
New client connections may be blocked or even fail.

Possible Causes

Too many connections have been established with the Impala server or the threshold is too small.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Connections** > **Number of ODBC Connections to Impalad Process (Impalad)** to check the threshold.

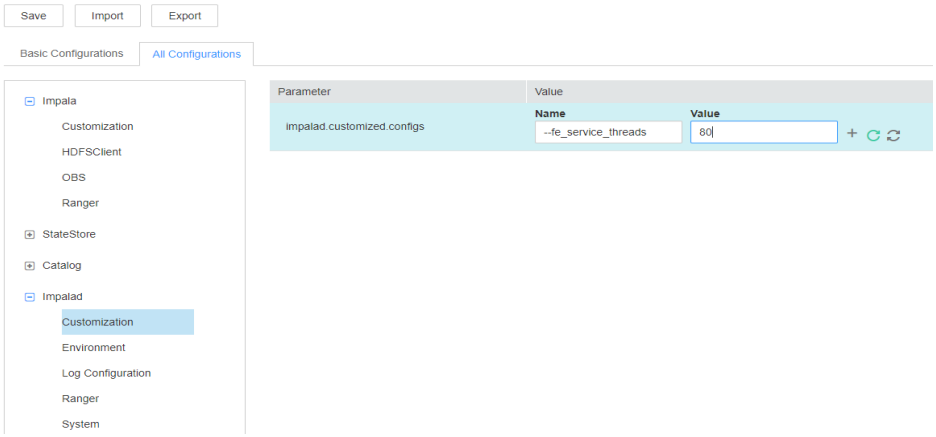


Step 2 Check the number of ODBC applications connected to Impalad and stop idle applications. Check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3** to change the number of concurrent connections supported by Impalad.

Step 3 On FusionInsight Manager, choose **Cluster** > **Impala** > **Configurations** > **All Configurations** > **Impalad** > **Customization**. Add the customized parameter **--fe_service_threads**. The default value of this parameter is **64**. Change the value as required and click **Save**.

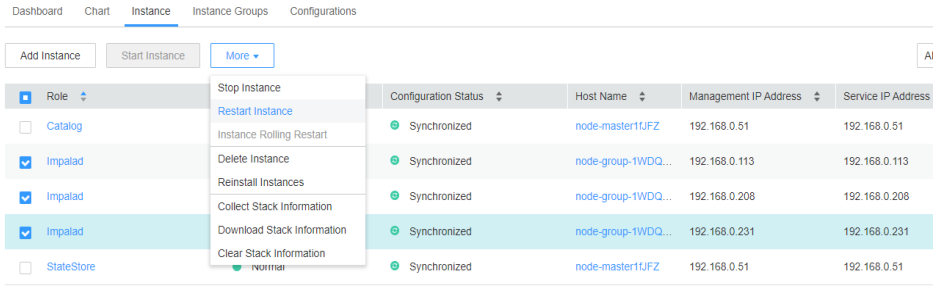
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



Step 4 After the query tasks on all clients are complete, click the **Instances** tab. Select all Impalad instances, and restart them.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.



Step 5 After the restart is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

After the fault that triggers the alarm is rectified, the alarm is automatically cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.246 ALM-29007 Impalad Process Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the Impalad process every 30 seconds. This alarm is generated when the system detects that the memory usage exceeds the default threshold (80%).

This alarm is automatically cleared when the system detects that the memory usage of the process falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29007	Minor	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The memory usage is too high. Some query tasks may fail due to insufficient memory.

Possible Causes

The Impalad process is executing a large number of query tasks.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **CPU and Memory** > **Impalad Process Memory Usage (Impalad)** and check the threshold.

Step 2 If the alarm threshold is smaller than 80%, increase the alarm threshold as required and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 If the threshold is greater than 80%, check whether a large number of concurrent query tasks exist when the alarm is generated. A large number of concurrent query tasks will cause the memory usage to increase sharply. After the tasks are complete, check whether the alarm is automatically cleared. During this period, some tasks may fail to be executed or may be canceled due to insufficient memory. In this case, try again.

NOTE

If the memory usage always exceeds the threshold, the cluster capacity needs to be expanded.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager of the active or standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

The alarm is automatically cleared after the burst concurrent tasks are complete.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.247 ALM-29008 Number of ODBC Connections to Impalad Exceeds the Threshold

Alarm Description

The system checks the number of client connections to the Impalad node every 30 seconds. This alarm is generated when the number of client connections exceeds the customized threshold (60 by default).

This alarm is automatically cleared when the number of client connections is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29008	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

New client connections may be blocked or even fail.

Possible Causes

Too many connections have been established with the Impala server or the threshold is too small.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Connections** > **Number of ODBC Connections to Impalad Process (Impalad)** to check the threshold.

Step 2 Check the number of ODBC applications connected to Impalad and stop idle applications. Check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 3** to change the number of concurrent connections supported by Impalad.

Step 3 On FusionInsight Manager, choose **Cluster** > **Impala** > **Configurations** > **All Configurations** > **Impalad** > **Customization**. Add the custom parameter **--fe_service_threads**. The default value of this parameter is **64**. Change the value as required and click **Save**.

Step 4 After the query tasks on all clients are complete, click the **Instances** tab. Select all Impalad instances, and restart them.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.

Step 5 After the restart is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If yes, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.248 ALM-29010 Number of Queries Being Submitted by Impalad Exceeds the Threshold

Alarm Description

The system checks the total number of queries being submitted by the Impalad node every 60 seconds. This alarm is generated when the number of queries exceeds the customized threshold (150 by default).

This alarm is automatically cleared when the number of queries is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29010	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The queries may be blocked or even fail.

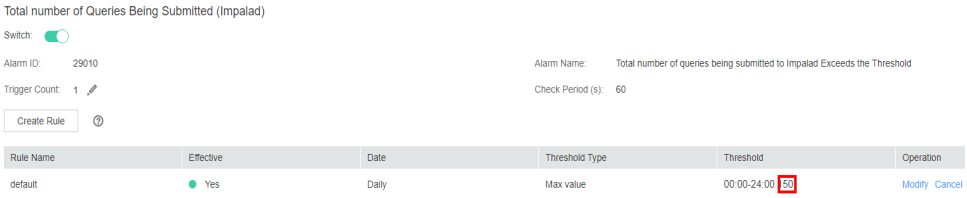
Possible Causes

The Impalad service has maintained a large number of queries, or the threshold is too small.

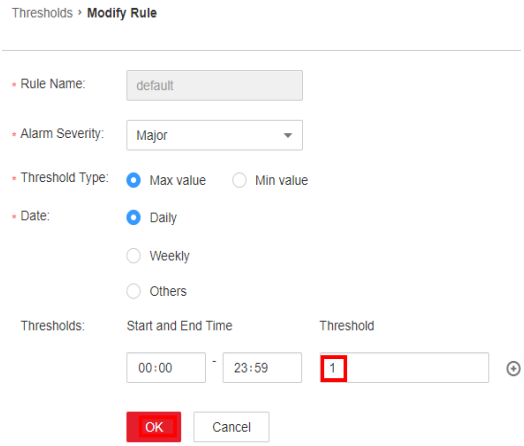
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Query Task Sum Statistics** > **Total number of Queries Being Submitted (Impalad)** and check the threshold.



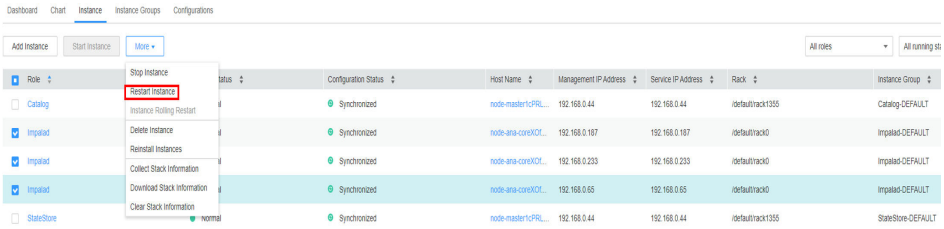
Step 2 Change the threshold.



Step 3 Click the **Instances** tab, select all Impalad instances, and restart them.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.



Step 4 After the restart is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.249 ALM-29011 Number of Queries Being Executed by Impalad Exceeds the Threshold

Alarm Description

The system checks the total number of queries being executed by the Impalad node every 60 seconds. This alarm is generated when the number of queries exceeds the customized threshold (150 by default).

This alarm is automatically cleared when the number of queries is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29011	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

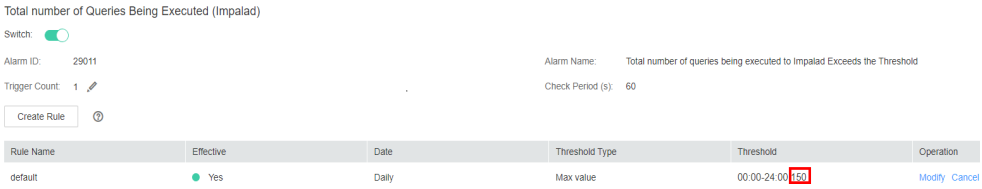
The queries may be blocked or even fail.

Possible Causes

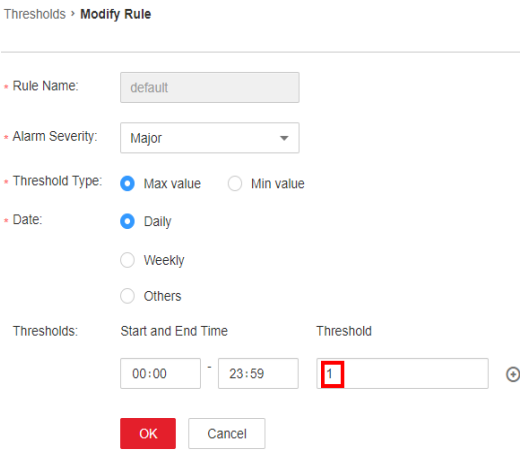
The Impalad service has maintained a large number of queries, or the threshold is too small.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Query Task Sum Statistics** > **Total number of Queries Being Executed (Impalad)** and check the threshold.



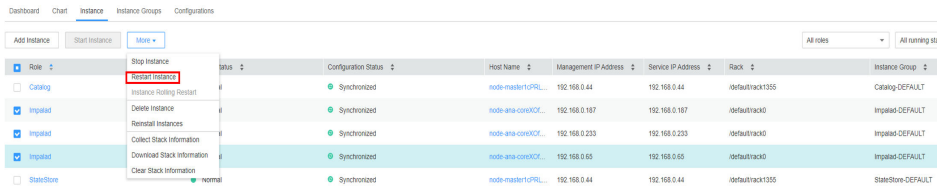
Step 2 Change the threshold.



Step 3 Click the **Instances** tab, select all Impalad instances, and restart them.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.



Step 4 After the restart is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.250 ALM-29012 Number of Queries Being Waited by Impalad Exceeds the Threshold

Alarm Description

The system checks the total number of queries being waited by the Impalad node every 60 seconds. This alarm is generated when the number of queries exceeds the customized threshold (150 by default).

This alarm is automatically cleared when the number of queries is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29012	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

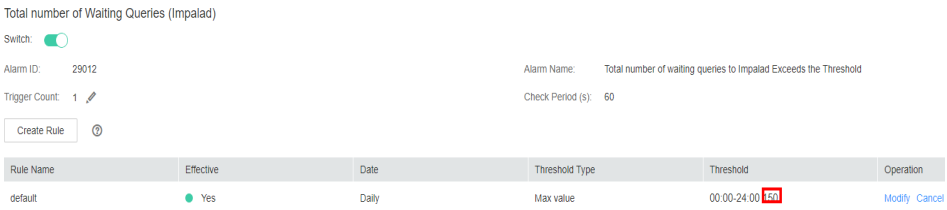
The queries may be blocked or even fail.

Possible Causes

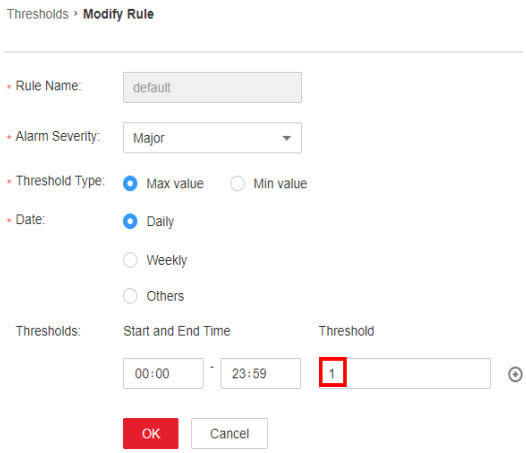
The Impalad service has maintained a large number of queries, or the threshold is too small.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Query Task Sum Statistics** > **Total number of Waiting Queries (Impalad)** and check the threshold.



Step 2 Change the threshold.

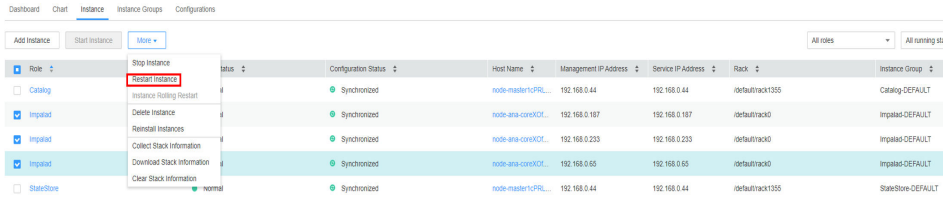


Step 3 Click the **Instances** tab, select all Impalad instances, and restart them.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.



Step 4 After the restart is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.251 ALM-29013 Impalad FGC Time Exceeds the Threshold

Alarm Description

The system checks the FGC time of the Impalad service every 60 seconds. This alarm is generated when the FGC time exceeds the threshold (12 seconds) for five consecutive times. This alarm is cleared when the FGC time is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29013	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Data read and write are affected.

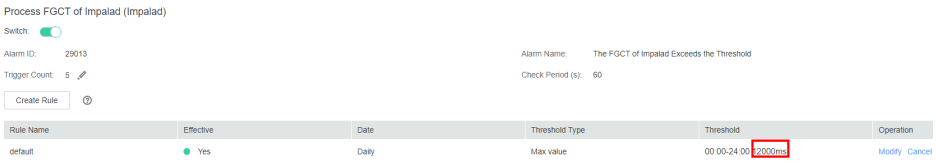
Possible Causes

The memory of the node instance is overused or the heap memory is inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC time.

Step 1 Choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Process FGCT** > **Process FGCT of Impalad (Impalad)**, and check the threshold (12s by default).



Step 2 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether the alarm whose **Alarm ID** is **29013** exists in the alarm list.

- If yes, go to **3**.
- If no, no further action is required.

Step 3 On FusionInsight Manager, choose **Cluster** > **Impala**, click the **Instances** tab, select the Impalad instance for which the alarm is generated, then click the **Chart** tab, locate the **Process FGCT** chart, and check whether the FGC time is greater than the threshold in **1**.

- If yes, go to **4**.
- If no, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **Process FGCT** > **Process FGCT of Impalad (Impalad)**, and change the threshold to a value less than the time obtained in **3**. Then, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager of the active or standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.252 ALM-29014 Catalog FGC Time Exceeds the Threshold

Alarm Description

The system checks the FGC time of the Catalog service every 60 seconds. This alarm is generated when the FGC time exceeds the threshold (12 seconds) for five consecutive times. This alarm is cleared when the FGC time is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29014	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Data read and write are affected.

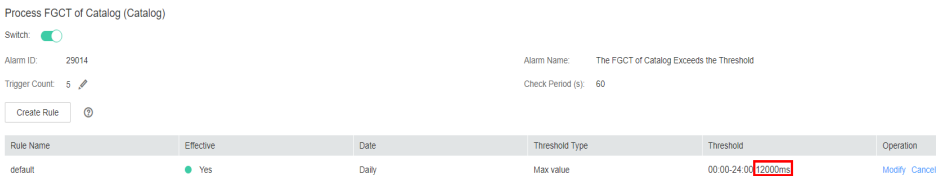
Possible Causes

The memory of the node instance is overused or the heap memory is inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC time.

Step 1 Choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, choose **Impala** > **Process FGCT** > **Process FGCT of Catalog (Catalog)**, and check the threshold (12s by default).



Step 2 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether the alarm whose **Alarm ID** is **29014** exists in the alarm list.

- If yes, go to **3**.
- If no, no further action is required.

Step 3 On FusionInsight Manager, choose **Cluster** > **Impala**, click the **Instance** tab, select the Catalog instance for which the alarm is generated, then click the **Chart** tab, locate the **Process FGCT** chart, and check whether the FGC time is greater than the threshold in **1**.

- If yes, go to **4**.
- If no, go to **5**.

Step 4 Choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, choose **Impala** > **Process FGCT** > **Process FGCT of Catalog (Catalog)**, and change the threshold to a value less than the time obtained in **3**. Then, check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to 5.

Collect fault information.

Step 5 On FusionInsight Manager of the active or standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.253 ALM-29015 Catalog Process Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the Catalog process every 30 seconds. This alarm is generated when the system detects that the memory usage exceeds the default threshold (80%).

This alarm is automatically cleared when the system detects that the memory usage of the process falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29015	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The memory usage is too high. Some query tasks may fail due to insufficient memory.

Possible Causes

The memory of the node instance is overused or the memory is inappropriately configured.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds** > **Impala** > **CPU and Memory** > **Catalog Process Memory Usage (Impalad)** and check the threshold.

Step 2 If the alarm threshold is smaller than 80%, increase the alarm threshold as required and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 If the threshold is greater than 80%, check whether a large number of concurrent query tasks exist when the alarm is generated. A large number of concurrent query tasks will cause the memory usage to increase sharply. After the tasks are complete, check whether the alarm is automatically cleared. During this period, some tasks may fail to be executed or may be canceled due to insufficient memory. In this case, try again.

NOTE

If the memory usage always exceeds the threshold, the cluster capacity needs to be expanded.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager of the active or standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.254 ALM-29016 Impalad Instance in the Sub-healthy State

Alarm Description

In MRS 3.1.5, the system checks every 60 seconds whether the Hive Server2 HTTP port (28000) of Impalad responds to cURL requests. This alarm is generated when the returned result has been incorrect for 20 seconds in two consecutive times. This alarm is cleared when the system correctly responds within 20 seconds.

In other MRS versions, the system checks every 60 seconds whether Impalad can execute **select 1**. This alarm is generated when the returned result has been incorrect for 20 seconds in two consecutive times. This alarm is cleared when the SQL statement is correctly executed within 20 seconds.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29016	Minor	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

If Impalad cannot execute SQL or experiences SQL execution timeouts, data read and write will be affected.

Possible Causes

The Impalad service maintains too many queries.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Impala** > **Impalad Web UI**. On the displayed page, click any node to go to the web UI.

Step 2 On the web UI, click **/backends** to view the Impala instance list. Locate the instance for which the alarm is generated and click **Web UI**. After the web UI of the subhealthy node is displayed, click **/queries** to check the task execution status and check whether any task is executed slowly.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 After the task is complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Impala** > **Instances**, select the Impala instance for which the alarm is generated, click **More**, and select **Restart Instance**. Then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

NOTE

The service will become unavailable when all instances are restarted. If a single instance is restarted, the tasks that are being executed on that instance will fail and the service will become available.

Collect fault information.

Step 5 On FusionInsight Manager of the active or standby cluster, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Impala** for the target cluster.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.255 ALM-29104 Tserver Process Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the Kudu Tserver process every 60 seconds. This alarm is generated when the system detects that the memory usage of the Kudu Tserver process exceeds the threshold.

This alarm is cleared when the memory usage of the Tserver process becomes normal and the system considers that the Kudu instance service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29104	Critical	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Users cannot use the Kudu service.

Possible Causes

The memory usage of the KuduTserver instance is too high.

Handling Procedure

Handle the Kudu instance exceptions.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed page, locate alarm **ALM-29104 Tserver Process Memory Usage Exceeds the Threshold** and view the alarm source.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > **Kudu**. Locate the threshold of this alarm and check whether the Kudu instance memory usage exceeds the threshold. If yes, rectify the fault or change the threshold.

Step 3 Choose **O&M** > **Alarm** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.256 ALM-29106 Tserver Process CPU Usage Exceeds the Threshold

Alarm Description

The system checks the Kudu service status every 60 seconds. This alarm is generated when the system detects that the CPU usage of the Kudu Tserver process is too high.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the CPU usage of the Tserver process becomes normal and the system considers that the Kudu instance service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29106	Critical	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Users cannot use the Kudu service.

Possible Causes

The CPU usage of a KuduTserver instance is too high.

Handling Procedure

Handle the Kudu instance exceptions.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed page, locate alarm **ALM-29106 Tserver Process CPU Usage Exceeds the Threshold** and view the alarm source.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > **Kudu**. Locate the alarm threshold and check whether the CPU usage of the cluster Kudu instance exceeds the threshold. If yes, rectify the fault or change the threshold.

Step 3 Choose **O&M** > **Alarm** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 6 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.257 ALM-29107 Tserver Process Memory Usage Exceeds the Threshold

Alarm Description

The system checks the Kudu service status every 60 seconds. This alarm is generated when the memory usage of the Kudu Tserver process exceeds the threshold.

This alarm is cleared when the memory usage of the Tserver process becomes normal and the system considers that the Kudu instance service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
29107	Critical	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Type	Parameter	Description
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Users cannot use the Kudu service.

Possible Causes

The memory usage of the KuduTserver instance is too high.

Handling Procedure

Handle the Kudu instance exceptions.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed page, locate alarm **ALM-29107 Tserver Process Memory Usage Exceeds the Threshold** and view the alarm source.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > **Kudu**. Locate the alarm threshold, compare the memory usage of the KuduTserver instance in the cluster with the threshold, and find the node whose memory usage exceeds the threshold.

Add nodes or reschedule jobs to reduce the memory usage of the Tserver node or change the threshold.

Step 3 Choose **O&M** > **Alarm** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.258 ALM-38000 Kafka Service Unavailable

Alarm Description

The system checks the Kafka service status every 30 seconds. This alarm is generated when the Kafka service is unavailable.

This alarm is cleared when the Kafka service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38000	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The cluster cannot provide the Kafka service, and users cannot perform new Kafka tasks.

Possible Causes

- The KrbServer service is faulty (for non-normal clusters).
- The ZooKeeper service is abnormal or does not respond.
- The Broker instances in the Kafka cluster are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the status of the KrbServer service. (Skip this step if the normal mode is used.)

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **KrbServer**.

Step 2 Check whether the running status of the KrbServer service is **Normal**.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

Step 3 Rectify the fault by following the steps provided in **ALM-25500 KrbServer Service Unavailable**.

Step 4 Perform **Step 2** again.

Check the status of the ZooKeeper cluster.

Step 5 Check whether the running status of the ZooKeeper service is **Normal**.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 If ZooKeeper service is stopped, start it. Otherwise, rectify the fault by following the steps provided in **ALM-13000 ZooKeeper Service Unavailable**.

Step 7 Perform **Step 5** again.

Check the Broker status.

Step 8 Choose **Cluster** > **Services** > **Kafka** and click **Instances**. The Kafka instance page is displayed.

Step 9 Check whether all instances in **Roles** are running properly.

- If yes, go to **Step 11**.
- If no, go to **Step 10**.

Step 10 Select all Broker instances, choose **More** > **Restart Instance**, and check whether the instances restart successfully.

⚠ CAUTION

During the restart of the Broker instance, if the current Topic is a single copy and is on the current Broker node, the Kafka service will be interrupted. Otherwise, the Kafka service will not be affected.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Choose **Cluster** > **Services** > **Kafka** and check whether the service status is **Normal**.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Wait for 30 seconds and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 14 Select **Kafka** in the required cluster from the **Service** drop-down list.

Step 15 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 16 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.259 ALM-38001 Insufficient Kafka Disk Capacity

Description

The system checks the Kafka disk usage every 60 seconds and compares the actual disk usage with the threshold. The disk usage has a default threshold. This alarm is generated when the disk usage is greater than the threshold.

You can change the threshold in **O&M** > **Alarm** > **Thresholds**. Under the service list, choose **Kafka > Disk > Broker Disk Usage (Broker)** and change the threshold.

When the **Trigger Count** is 1, this alarm is cleared when the Kafka disk usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the Kafka disk usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38001	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the disk partition where the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

Kafka data write operations are affected.

Possible Causes

- The configuration (such as number and size) of the disks for storing Kafka data cannot meet the requirement of the current service traffic, due to which the disk usage reaches the upper limit.
- Data retention time is too long, due to which the data disk usage reaches the upper limit.
- The service plan does not distribute data evenly, due to which the usage of some disks reaches the upper limit.

Procedure

Check the disk configuration of Kafka data.

Step 1 On the FusionInsight Manager portal and click **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, locate the alarm and obtain **HostName** from **Location**.

Step 3 Click **Cluster** > *Name of the desired cluster* > **Hosts**.

Step 4 In the host list, click the host name obtained in **Step 2**.

Step 5 Check whether the **Disk** area contains the partition name in the alarm.

- If yes, go to **Step 6**.
- If no, manually clear the alarm and no further operation is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Check whether the disk partition usage contained in the alarm reaches 100% in the **Disk** area.

- If yes, handle the alarm by following the instructions in **Related Information**.
- If no, go to **Step 7**.

Check the Kafka data storage duration.

Step 7 Choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Configurations**.

Step 8 Check whether the value of parameter **disk.adapter.enable** is set to **true**.

- If yes, go to **Step 10**.
- If no, go to **Step 9**.

Step 9 Set the value of **disk.adapter.enable** to **true**. Check whether the value of **adapter.topic.min.retention.hours** is properly set.

- If yes, go to **Step 10**.
- If no, adjust the data retention period based on service requirements.

NOTICE

If the disk auto-adaptation function is enabled, some historical data of specified topics is deleted. If the retention period of some topics cannot be adjusted, click **All Configurations** and add the topics to the value of the **disk.adapter.topic.blacklist** parameter.

Step 10 Wait 10 minutes and check whether the usage of faulty disks reduces.

- If yes, wait until the alarm is cleared.
- If no, go to **Step 11**.

Check the Kafka data plan.

Step 11 In the **Instance** area, click **Broker**. In the **Real Time** area of Broker, Click the drop-down menu in the Chart area and choose **Customize** to customize monitoring items.

Step 12 In the dialog box, select **Disk** > **Broker Disk Usage** and click **OK**.

The Kafka disk usage information is displayed.

Step 13 View the information in **Step 12** to check whether there is only the disk parathion for which the alarm is generated in **Step 2**.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Perform disk planning and mount a new disk again. Go to the **Instance Configurations** page of the node for which the alarm is generated, modify **log.dirs**, add other disk directories, and restart the Kafka instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

If the current topic has multiple replicas during the instance rolling restart, there is no impact on the Kafka service. Otherwise, the Kafka service will be unavailable during the restart, and upper-layer services that depend on the service will be affected.

Step 15 Determine whether to shorten the data retention time configured on Kafka based on service requirements and service traffic.

- If yes, go to **Step 16**.
- If no, go to **Step 17**.

Step 16 Log in to FusionInsight Manager, select **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Configurations**, and click **All Configurations**. In the search box on the right, enter **log.retention.hours**. The value of the parameter indicates the default data retention time of the topic. You can change the value to a smaller one.

NOTE

- For a topic whose data retention time is configured alone, the modification of the data retention time on the Kafka Service Configuration page does not take effect.
- To modify the data retention time for a topic, use the Kafka client command-line interface (CLI) to configure the topic.

  Example: **kafka-topics.sh --zookeeper "***ZooKeeper IP address***:2181/kafka" --alter --topic "***Topic bane***" --config retention.ms= "***retention time***"**

Step 17 Check whether the usage of some disks reaches the upper limit due to unreasonable configuration of the partitions of some topics. For example, the number of partitions configured for a topic with large data volume is smaller than the number of disks. In this case, the data is not evenly allocated to disks.

NOTE

If you do not know which topics have a large amount of service data, perform the following steps:

1. Log in to an instance node based on the host node information obtained in **Step 2**.
2. Go to the data directory (directory specified by **log.dirs** before the modification in **Step 14**).
3. Run the following command to check whether there is topic with partition that use large disk space.

   du -h --max-depth=1 ./

- If yes, go to **Step 18**.
- If no, go to **Step 19**.

Step 18 In the Kafka client CLI, run the following command to perform partition capacity expansion for the topic:

kafka-topics.sh --zookeeper "*ZooKeeper IP address***:2181/kafka" --alter --topic "***Topic name***" --partitions="***New number of partitions***"**

NOTE

- You are advised to set the new number of partitions to a multiple of the number of Kafka data disks.
- The step may not quickly clear the alarm, and you need to modify the data retention time in **Step 11** to gradually balance data allocation.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 19 Determine whether to perform capacity expansion.

NOTE

You are advised to perform capacity expansion for Kafka when the current disk usage exceeds 80%.

- If yes, go to **Step 20**.
- If no, go to **Step 21**.

Step 20 Expand the disk capacity and check whether the alarm is cleared after capacity expansion.

- If yes, no further action is required.
- If no, go to **Step 22**.

Step 21 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 22**.

Collect fault information.

Step 22 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 23 Select **Kafka** in the required cluster from the **Service** drop-down list.

Step 24 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 25 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

Step 1 Log in to FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance**, stop the Broker instance whose status is **Restoring**, record the management IP address of the node where the Broker instance is located, and record **broker.id**. The value can be obtained by using the following method: Click the role name. On the **Configurations** page, select **All Configurations**, and search for the **broker.id** parameter.

Step 2 Log in to the recorded management IP address as user **root**, and run the **df -lh** command to view the mounted directory whose disk usage is 100%, for example, **${BIGDATA_DATA_HOME}/kafka/data1**.

Step 3 Go to the directory and run the **du -sh *** command to view the size of each file in the directory. Check whether there are files in addition to the files in the **kafka-logs** directory, and determine whether these files can be deleted or migrated.

- If yes, go to **Step 8**.
- If no, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Go to the **kafka-logs** directory, run the **du -sh *** command, select a partition folder to be moved. The naming rule is **Topic name-Partition ID**. Record the topic and partition.

Step 5 Modify the **recovery-point-offset-checkpoint** and **replication-offset-checkpoint** files in the **kafka-logs** directory in the same way.

1. Decrease the number in the second line in the file. (To remove multiple directories, the number deducted is equal to the number of files to be removed.)
2. Delete the line of the to-be-removed partition. (The line structure is "Topic name Partition ID Offset". Save the data before deletion. Subsequently, the content must be added to the file of the same name in the destination directory.)

Step 6 Modify the **recovery-point-offset-checkpoint** and **replication-offset-checkpoint** files in the destination data directory. For example, **${BIGDATA_DATA_HOME}/kafka/data2/kafka-logs** in the same way.

- Increase the number in the second line in the file. (To move multiple directories, the number added is equal to the number of files to be moved.)
- Add the to-be moved partition to the end of the file. (The line structure is "Topic name Partition ID Offset". You can copy the line data saved in **Step 5**.)

Step 7 Move the partition to the destination directory. After the partition is moved, run the **chown omm:wheel -R** *Partition directory* command to modify the directory owner group for the partition.

Step 8 Log in to FusionInsight Manager and choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance** to start the Broker instance.

Step 9 Wait for 5 to 10 minutes and check whether the health status of the Broker instance is **Normal**.

- If yes, resolve the disk capacity insufficiency problem according to the handling method of "ALM-38001 Insufficient Kafka Disk Space" after the alarm is cleared.
- If no, contact the O&M personnel.

----End

9.11.260 ALM-38002 Kafka Heap Memory Usage Exceeds the Threshold

Description

The system checks the Kafka service status every 60 seconds. The alarm is generated when the heap memory usage of a Kafka instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times.

When the **Trigger Count** is 1, this alarm is cleared when the heap memory usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the heap memory usage is less than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38002	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available direct memory of the Kafka service is insufficient, a memory overflow occurs and the broker instance breaks down. As a result, the broker cannot provide read and write services properly.

Possible Causes

The heap memory of the Kafka instance is overused or the heap memory is inappropriately allocated.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **Kafka Heap Memory Usage Exceeds the Threshold** > **Location**. Check the host name of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Process** > **Heap Memory Usage of Kafka**, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the used heap memory of Kafka reaches 95% of the maximum heap memory specified for Kafka.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Check the heap memory size configured for Kafka.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Configurations** > **All Configurations**> **Broker(Role)** > **Environment**. Increase the value of **KAFKA_HEAP_OPTS** by referring to the Note.

NOTE

- It is recommended that **-Xmx** and **-Xms** be set to the same value.
- You are advised to view **Heap Memory Usage of Kafka** by referring to **Step 2**, and set the value of **KAFKA_HEAP_OPTS** to twice the value of **Heap Memory Used by Kafka.**

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 7 Select **Kafka** in the required cluster from the **Service** drop-down list.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.261 ALM-38004 Kafka Direct Memory Usage Exceeds the Threshold

Description

The system checks the direct memory usage of the Kafka service every 30 seconds. This alarm is generated when the direct memory usage of a Kafka instance exceeds the threshold (80% of the maximum memory) for 10 consecutive times.

When the **Trigger Count** is 1, this alarm is cleared when the direct memory usage is less than or equal to the threshold. When the **Trigger Count** is greater than 1,

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

this alarm is cleared when the direct memory usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38004	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the available direct memory of the Kafka service is insufficient, a memory overflow occurs and the broker instance breaks down. As a result, the broker cannot provide read and write services properly.

Possible Causes

The direct memory of the Kafka instance is overused or the direct memory is inappropriately allocated.

Procedure

Check the direct memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **ALM-38004 Kafka Direct Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services > Kafka > Instance. Click the instance for which the alarm is**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

generated to go to the page for the instance. Click the drop-down menu in the **Chart** area, choose **Customize** > **Process** > **Kafka Direct Memory Usage**, and click **OK**.

Step 3 Check whether the used direct memory of Kafka reaches 80% of the maximum direct memory specified for Kafka.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Check the direct memory size configured for the Kafka.

Step 4 On the FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka**. On the displayed page, click **Configurations** and then choose **All Configurations** > **Broker(Role)** > **Environment**. Increase the value of **-Xmx** in the **KAFKA_HEAP_OPTS** parameter.

NOTE

- It is recommended that **-Xmx** and **-Xms** be set to the same value.
- You are advised to view **Kafka Direct Memory Usage** by referring to **Step 2**, and set the value of **KAFKA_HEAP_OPTS** to twice the value of **Direct Memory Used by Kafka.**

Step 5 Save the configuration and restart the Kafka service.

NOTE

If rolling restart is performed and the current Topic has multiple copies, services are not affected. Otherwise, the Kafka service will be unavailable during the restart, and upper-layer services that depend on the Kafka service will be affected.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Kafka** in the required cluster from the **Service** drop-down list.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.262 ALM-38005 GC Duration of the Broker Process Exceeds the Threshold

Description

The system checks the garbage collection (GC) duration of the Broker process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default) for 3 consecutive times.

When the **Trigger Count** is 1, this alarm is cleared when the GC duration is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the GC duration is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38005	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

A long GC duration of the Broker process may interrupt the services.

Possible Causes

The Kafka GC duration of the node is too long or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the GC duration.

Step 1 On the FusionInsight Manager portal, choose **O&M** > **Alarm** > **Alarms** > **GC Duration of the Broker Process Exceeds the Threshold** > **Location**. Check the host name of the instance involved in this alarm.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > **Services** > **Kafka** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Process** > **Broker GC Duration per Minute**, and click **OK**.

Step 3 Check whether the GC duration of the Broker process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Check the heap memory size configured for the Kafka.

Step 4 On the FusionInsight Manager, choose **Cluster** > **Services** > **Kafka**. On the displayed page, click **Configurations** and then choose **All Configurations** > **Broker(Role)** > **Environment**. Increase the value of **-Xmx** in the **KAFKA_HEAP_OPTS** parameter.

NOTE

- It is recommended that **-Xmx** and **-Xms** be set to the same value.
- You are advised to set the value of **KAFKA_HEAP_OPTS** to twice the value of **Heap Memory Utilization of Kafka**. On the FusionInsight Manager portal, choose **Cluster** > **Services** > **Kafka** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **Process** > **Heap Memory Utilization of Kafka** to check the value of heap memory used by Kafka.

Step 5 Save the configuration and restart the Kafka service.

NOTE

If rolling restart is performed and the current Topic has multiple copies, services are not affected. Otherwise, the Kafka service will be unavailable during the restart, and upper-layer services that depend on the Kafka service will be affected.

Step 6 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Kafka** in the required cluster from the **Service** drop-down list.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.263 ALM-38006 Percentage of Kafka Partitions That Are Not Completely Synchronized Exceeds the Threshold

Description

The system checks the percentage of Kafka partitions that are not completely synchronized to the total number of partitions every 60 seconds. This alarm is generated when the percentage exceeds the threshold (50% by default) for 3 consecutive times.

When the **Trigger Count** is 1, this alarm is cleared when the percentage is less than or equal to the threshold. When the **Trigger Count** is greater than 1, this alarm is cleared when the percentage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38006	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Too many Kafka partitions that are not completely synchronized affect service reliability. In addition, data may be lost when leaders are switched.

Possible Causes

Some nodes where the Broker instance resides are abnormal or stop running. As a result, replicas of some partitions in Kafka are out of the in-sync replicas (ISR) set.

Procedure

Check Broker instances.

Step 1 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance**. The Kafka instances page is displayed.

Step 2 Check whether faulty nodes exist among all Broker nodes.

- If yes, record the host name of the node and go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Check whether any fault alarm is generated for the host in **Step 2**. If yes, clear the alarm based on the alarm guide.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance**. The Kafka instances page is displayed.

Step 5 Check whether stopped nodes exist among all Broker instance.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 Select all stopped Broker instances and click **Start Instance**.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 9 Select **Kafka** in the required cluster from the **Service** drop-down list.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.264 ALM-38007 Status of Kafka Default User Is Abnormal

Description

The system checks the default user of Kafka every 60 seconds. This alarm is generated when the system detects that the user status is abnormal.

Trigger Count is set to **1**. This alarm is cleared when the user status becomes normal.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38007	Critical	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host name for which the alarm is generated.
Trigger Condition	Specifies the condition that the Kafka default user status is abnormal.

Impact on the System

If the Kafka default user status is abnormal, metadata synchronization between Brokers and interaction between Kafka and ZooKeeper will be affected, affecting service production, consumption, and topic creation and deletion.

Possible Causes

- The Sssd service is abnormal.
- Some Broker instances stop running.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check whether the Sssd service is abnormal.

Step 1 On the FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **Status of Kafka Default User Is Abnormal** > **Location** to check the host name of the instance for which the alarm is generated.

Step 2 Find the host information in the alarm information and log in to the host.

Step 3 Run the **id -Gn kafka** command and check whether "No such user" is displayed in the command output.

- If yes, record the host name of the node and go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Alarms**. Check whether there is **Sssd Service Exception** in the alarm information. If there is, handle the alarm based on alarm information.

Check the running status of the Broker instance.

Step 5 On the FusionInsight Manager home page, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance**. The Kafka instance page is displayed.

Step 6 Check whether there are stopped nodes on all Broker instances.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Select all stopped Broker instances and click **Start Instance**.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 10 In the **Service** area, select **Kafka** in the required cluster.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.265 ALM-38008 Abnormal Kafka Data Directory Status

Description

The system checks the Kafka data directory status every 60 seconds. This alarm is generated when the system detects that the status of a data directory is abnormal.

Trigger Count is set to **1**. This alarm is cleared when the data directory status becomes normal.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38008	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host name for which the alarm is generated.
DirName	Specifies the directory name for which the alarm is generated.
Trigger Condition	Specifies the condition that the Kafka data directory status is abnormal.

Impact on the System

If the Kafka data directory status is abnormal, the current replicas of all partitions in the data directory are brought offline, and the data directory status of multiple nodes is abnormal at the same time. As a result, some partitions may become unavailable.

Possible Causes

- The data directory permission is tampered with.
- The disk where the data directory is located is faulty.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the permission on the faulty data directory.

Step 1 Find the host information in the alarm information and log in to the host.

Step 2 In the alarm information, check whether the data directory and its subdirectories belong to the **omm:wheel** group.

- If yes, record the host name of the node and go to **Step 4**.
- If no, go to **Step 3**.

Step 3 Restore the owner group of the data directory and its subdirectories to **omm:wheel**.

Check whether the disk where the data directory is located is faulty.

Step 4 In the upper-level directory of the data directory, create and delete files as user **omm**. Check whether data read/write on the disk is normal.

- If yes, go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Replace or repair the disk where the data directory is located to ensure that data read/write on the disk is normal.

Step 6 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Kafka** > **Instance**. On the Kafka instance page that is displayed, restart the Broker instance on the host recorded in **Step 2**.

NOTE

During the restart of the Broker instance, if the current Topic is a single copy and is on the current Broker node, the Kafka service will be interrupted. Otherwise, the Kafka service will not be affected.

Step 7 After Broker is started, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 9 In the **Service** area, select **Kafka** in the required cluster.

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.266 ALM-38009 Busy Broker Disk I/Os (Applicable to Versions Later Than MRS 3.1.0)

NOTE

This section applies to versions later than MRS 3.1.0.

Alarm Description

The system checks the I/O status of each Kafka disk every 60 seconds. This alarm is generated when the disk I/O of a Kafka data directory on a broker exceeds the threshold (80% by default).

Its **Trigger Count** is **3**. This alarm is cleared when the disk I/O is lower than the threshold (80% by default).

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38009	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
	DataDirectory Name	Specifies the name of the Kafka data directory with frequent disk I/Os.

Impact on the System

The disk partition has frequent I/Os. Data may fail to be written to the Kafka topic for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- There are many replicas configured for the topic.
- The parameter for batch writing producer's messages is inappropriately configured. The service traffic of this topic is too heavy, and the current partition configuration is inappropriate.

Handling Procedure

Check the number of topic replicas.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the host name for which this alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Kafka** > **KafkaTopic Monitor**, and sort topics by the number of replicas in descending order. Check whether topics with more than three replicas are experiencing high traffic.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 If such topics are found, click the topic name and verify whether the IP address resolved from the host name obtained in **Step 1** is included among the hosts for all partitions of the topic. If it is, reduce the topic's replication factor to 3.

Perform the operations below on the Kafka client to replan replicas for Kafka topics.

1. Log in to the node where the Kafka client is installed as the client installation user.
2. Run the following command to switch to the client installation directory, for example, **/opt/client/Kafka/kafka/bin**:
   ```
   cd /opt/client/Kafka/kafka/bin
   ```
3. Run the following command to configure environment variables:
   ```
   source /opt/client/bigdata_env
   ```
4. Run the following command to perform user authentication (skip this step for a cluster in normal mode):
   ```
   kinit Component service user
   ```
5. Run the following command to replan the replicas of Kafka topics:
   ```
   kafka-reassign-partitions.sh --bootstrap-server <Kafka cluster IP address:21007> --command-config ../config/client.properties --reassignment-json-file {manual assignment json file path} --execute
   ```

   To obtain the IP address of the Kafka cluster, log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kafka** > **Instances**. Check and record the service IP address of any Broker instance. The port number of the Kafka cluster is 21007 if Kerberos authentication is enabled for the cluster (the cluster is in security mode) and is 21005 if Kerberos authentication is disabled for the cluster (the cluster is in normal mode).

   Example:

   ```
   /opt/client/Kafka/kafka/bin/kafka-reassign-partitions.sh --bootstrap-server 192.168.0.90:21007,192.168.0.91:21007,192.168.0.92:21007 --command-config /opt/client/Kafka/kafka/config/client.properties --reassignment-json-file expand-cluster-reassignment.json --execute
   ```

   In the **expand-cluster-reassignment.json** file, specify the target brokers to which the topic's partitions will be migrated. The JSON content should follow this format:

   ```
   {"partitions":[{"topic": "topicName","partition": 1,"replicas": [1,2,3] }],"version":1}.
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Observe for a period of time and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the partition planning of the topic.

Step 5 On the **KafkaTopic Monitor** page, view **Topic Input Traffic** in the **Topic Traffic** area of each topic, obtain the topic with the largest value, and check the partitions of this topic as well as information about the host of these partitions.

Step 6 Log in to the host queried in **Step 5** and run the following command to check the **%util** value of each disk:

```
iostat -d -x
```

```
:/opt/R3/FusionInsight_Manager/software/packs # iostat -d -x
Linux 3.0.76-0.11-default (189-39-172-162)      06/26/19        _x86_64_

Device:         rrqm/s   wrqm/s     r/s     w/s    rsec/s    wsec/s avgrq-sz avgqu-sz   await  svctm  %util
xvda              0.04    44.44    1.26   21.94     43.62    531.02    24.78     0.03    1.44   0.56   1.30
xvde              0.16   431.84   13.78   82.51    284.32   4115.90    45.70     0.06    1.41   0.64   6.21
```

- If the **%util** value of each disk exceeds the threshold (**80%** by default), expand the Kafka disk capacity. After the capacity expansion, replan the topic partitions by referring to **Step 3**.
- If the **%util** values of the disks vary greatly, check the disk partition configuration of Kafka.

  For example, check the value of **log.dirs** in the **${BIGDATA_HOME}/FusionInsight_HD_*/x_x_Broker/etc/server.properties** file.

  Run the following command to view the **Filesystem** information:

  ```
  df -h log.dirs value
  ```

  The command output is as follows.

  ```
  :/opt/R3/FusionInsight_Manager/software/packs # df -h /srv/BigData/kafka/data1/kafka-logs/
  Filesystem      Size  Used Avail Use% Mounted on
  /dev/xvda2       36G   21G   14G  62% /
  :/opt/R3/FusionInsight_Manager/software/packs #
  ```

- If the partition where Filesystem is located matches the partition with a high **%util** value, plan Kafka partitions on idle disks, configure **log.dirs** as an idle disk directory, and replan topic partitions by referring to **Step 3**. Ensure that the partitions of the topic are evenly distributed to each disk.

Step 7 Observe for a period of time and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, repeat **Step 5** to **Step 6** three times. Then, go to **Step 8**.

Step 8 Observe for a period of time and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 11 Click the edit icon in the upper right corner and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.267 ALM-38010 Topics with Single Replica

Description

The system checks the number of replicas of each topic every 60 seconds on the node where the Kafka Controller resides. This alarm is generated when there is one replica for a topic.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38010	Major	No

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
TopicName	Specifies the list of topics for which the alarm is generated.

Impact on the System

There is the single point of failure (SPOF) risk for topics with only one replica. When the node where the replica resides becomes abnormal, the partition does not have a leader, and services on the topic are affected.

Possible Causes

The number of replicas for the topic is incorrectly configured.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the number of replicas for the topic.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, click ⌄ of this alarm, and view the **TopicName** list in **Location**.

Step 2 Check whether replicas need to be added for the topic for which the alarm is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 On the FusionInsight client, re-plan topic replicas and describe the partition distribution of the topic in the **add-replicas-reassignment.json** file in the following format: {"partitions":[{"topic": "*topic name*","partition": 1,"replicas": [1,2] }],"version":1}. Then, run the following command to add replicas:

kafka-reassign-partitions.sh --zookeeper *{zk_host}:{port}*/**kafka --reassignment-json-file** *{manual assignment json file path}* **--execute**

For example:

/opt/client/Kafka/kafka/bin/kafka-reassign-partitions.sh --zookeeper 192.168.0.90:2181,192.168.0.91:2181,192.168.0.92:2181/kafka --reassignment-json-file add-replicas-reassignment.json --execute

Step 4 Run the following command to check the task execution progress:

kafka-reassign-partitions.sh --zookeeper *{zk_host}:{port}*/**kafka --reassignment-json-file** *{manual assignment json file path}* **--verify**

For example:

/opt/client/Kafka/kafka/bin/kafka-reassign-partitions.sh --zookeeper 192.168.0.90:2181,192.168.0.91:2181,192.168.0.92:2181/kafka --reassignment-json-file add-replicas-reassignment.json --verify

Step 5 After completing the handling operations or confirming that the alarm has no impact, manually clear the alarm on FusionInsight Manager.

Step 6 After a period of time, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 8 In the **Service** area, select **Kafka** in the required cluster.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

If the alarm has no impact, manually clear the alarm.

Related Information

None

9.11.268 ALM-38011 User Connection Usage on Broker Exceeds the Threshold

Description

The system checks the number of connections of each user on Broker every 30 seconds. This alarm is generated when the connection usage of a user on the Broker exceeds the threshold (80% by default) for five consecutive times.

The number of times that smoothing is performed is **5**. This alarm is cleared when the connection usage of a user on the Broker is less than the threshold.

The alarm can be automatically cleared. However, if the number of connections of a user suddenly becomes **0** and no connection is created, the alarm cannot be automatically cleared. In this case, you need to manually clear it.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
38011	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
UserName	Specifies the username for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the number of connections of a user is excessive, the user cannot create new connections to the Broker.

Possible Causes

- The number of connections (created by a user) used by the client exceeds the preset threshold.
- The threshold for the connection usage does not meet service requirements.

Handling Procedure

Check the number of connections established by the same user on the client.

Step 1 On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Alarms** > **User Connection Usage on Broker Exceeds the Threshold**. Check the host name and username of the Broker instance for which the alarm is generated in **Location**.

Step 2 On the FusionInsight Manager portal, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Instance**. Click the instance for which the alarm is generated to go to the page for the instance. Click the drop-down list in the upper right corner of the icon area, choose **Customize** > **Other**, and select **User Connection Usage on Broker**, **Maximum Number of User Connections on Broker**, and **Number of User Connections on Broker** to view the number of the current user connections on the Broker.

Step 3 Observe the number of real-time connections of the current alarm user and check whether the real-time monitoring data of the current user exists.

- If yes, go to **Step 4**.
- If no, the current user has disconnected all connections. You need to clear the alarm manually, and no further action is required.

  NOTE

  After the alarm user disconnects all connections, the monitoring data of the user disappears. In this case, the alarm will not be automatically cleared. You need to manually clear it.

Step 4 Check whether the user is authorized by the service side.

- If yes, go to **Step 7**.
- If no, go to **Step 5**.

Step 5 Run the following command on the client to limit the number of connections of the user. There are two configuration rules based on the following commands:

1. For the specific Broker and user, run the following command:

   kafka-configs.sh --bootstrap-server *<broker ip:port>* **--alter --add-config 'max.connections.per.user.overrides=[***<username>***:***<connection.number>***]' --entity-type brokers --entity-name** *<broker.id>* **--command-config Kafka/kafka/config/producer.properties**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

For unauthorized users, confirm with the service side to reduce the maximum number of connections of an unauthorized user or set the maximum number of connections to **0**.

In the command, you need to specify the IP address and port number of Broker, set values of configuration items, and specify the **brokerId** and **username**. Here, the user refers to the authorized Kerberos user.

The configuration updated using the command line tool takes effect dynamically. The configuration becomes invalid after the service is restarted. To make the configuration take effect after the restart, log in to FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Configurations** > **All Configurations** > **Broker** > **Server**, and update the configuration to **max.connections.per.user.overrides**.

2. For the specific use and default Broker (that is, all Broker instances in the cluster), run the following command:

   kafka-configs.sh --bootstrap-server *<broker ip:port>* **--alter --add-config 'max.connections.per.user.overrides=[***<username>***:***<connection.number>***]' --entity-type brokers ---entity-default --command-config Kafka/kafka/config/client.properties**

   Example:

   kafka-configs.sh --bootstrap-server 10.153.3.26:21007 --alter --add-config 'max.connections.per.user.overrides=[showcase:4]' --entity-type brokers --entity-name 1 --command-config Kafka/kafka/config/client.properties

Step 6 Check whether the maximum number of connections is **0** and whether the number of connections of the current user decreases or remains unchanged according to **Step 2**.

- If yes, manually clear the alarm and no further action is required.
- If no, go to **Step 7**.

Step 7 Check whether the number of real-time connections and connection usage of the current user are sharply increased when they are compared with historical data, and whether have exceeded the specified maximum number of connections.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

NOTE

If there is an obvious increase after the comparison and the maximum number of connections has reached the preset value, the connections of the user may be abnormal. You need to confirm with the service party.

Check whether the number of user connections meets service requirements.

Step 8 Check whether the number of connections of the user meets service requirements.

- If yes, go to **Step 9**.
- If no, contact the service party to rectify the fault.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

If the number of user connections is abnormal, contact the service party to rectify the fault from the following aspects:

- Check whether new services are added so that the number of user connections increases sharply.
- Check whether handle leakage occurs on the code at the service side.

Step 9 Consider whether to increase the maximum number of connections of the user.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Increase the maximum number of connections based on the service requirements. Set the number of connections of the user on the Kafka client. For details, see **Step 5**.

Step 11 Wait for several minutes and then check whether the alarm is automatically cleared.

- If yes, go to **Step 12**.
- If no, go to **Step 2**.

Step 12 Determine whether to add the user to the whitelist based on service requirements on the service side.

- If yes, go to **Step 13**.
- If no, go to **Step 15**.

NOTE

To add a user to the whitelist, you need to restart the Kafka service. However, this operation will cause service interruption and affect service running. Therefore, you must confirm with the service side before performing this operation.

Step 13 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Kafka** > **Configurations** > **All Configurations** > **Broker(Role)** > **Server**, and add the user to the **max.connections.per.user.whitelist** configuration item.

Step 14 Restart the service for the modification to take effect. In addition, you need to manually clear the alarm, and no further action is required.

Collect the fault information.

Step 15 On the FusionInsight Manager homepage, choose **O&M** > **Log** > **Download**.

Step 16 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 17 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 18 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.269 ALM-38012 Number of Broker Partitions Exceeds the Threshold

Alarm Description

The system checks the number of partitions on each Broker instance of the Kafka service every 30 seconds. You can view the number of partitions on the Broker instance dashboard page. This alarm is generated when the number of partitions on a Broker instance exceeds the threshold. You can choose **O&M** > **Alarm** > **Thresholds** > **Kafka** and change the threshold. This alarm is cleared when the number of partitions is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38012	Critical (default threshold: 6000) Major (default threshold: 3000)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

The number of Broker partitions exceeds the threshold. Too many partitions increase the Broker load and cause bottlenecks in memory, disk I/O, and CPU resources. As a result, the request response becomes slow or even times out.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- Broker partitions are unevenly distributed, or the Kafka cluster usage exceeds the specifications.
- There are many useless topics.

Handling Procedure

Check whether partitions are evenly distributed on Broker.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the service instance and host for which the alarm is generated.

Step 2 Choose **Cluster** > **Services** > **Kafka** > **Chart**, select **Partition** from the **Chart Category** area, zoom in **Number of Partitions-All Instances** in the upper right corner, and click **Distribution** to check whether partitions are evenly distributed on Broker.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 If the partitions on Broker are balanced, the Kafka cluster exceeds the specifications. In this case, add Broker instances. Then, go to **Step 5**.

On FusionInsight Manager, choose **Cluster** > **Services** > **Kafka** > **Instances**, click **Add Instance**, and add Broker instances as prompted.

Step 4 Click the uneven distribution bar on the rightmost. If the number of partitions on only the Broker node for which the alarm is generated is too large, perform data balancing.

Step 5 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether there are many useless topics.

Step 6 Check whether useless topics exist in the cluster.

- If yes, perform the following steps to delete useless topics: **Deleting topics is a high-risk operation. Before deleting topics, ensure that the topics are not used.**
  a. Log in to Manager as a user who has the permission to access the Kafka web UI.
  b. Choose **Cluster** > **Services** > **Kafka**. On the right of **KafkaManager Web UI**, click the URL to access the Kafka web UI.
  c. Choose **Topics**.
  d. In the **Operation** column of the target topic, click **Action** and select **Delete**.
  e. In the dialog box that is displayed, click **OK**.
     The default built-in topics cannot be deleted.
- If no, go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.270 ALM-38013 Produce Request Latency in the Request Queue Exceeds the Threshold

Alarm Description

The system checks the latency of Produce requests on Broker instances in the request queue every 30 seconds. This alarm is generated when the latency of Produce requests on a Broker instance in the request queue has exceeded the threshold for 10 consecutive times.

This alarm is cleared when the latency of Produce requests in the request queue is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38013	Critical (default threshold: 60000) Major (default threshold: 30000)	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

The latency of Produce requests on the Broker instance in the request queue exceeds the threshold. As a result, the request queue is congested, and the response time of write requests increases. For latency-sensitive services, a large number of write requests may time out.

Possible Causes

- The number of threads used by the Broker instance to process requests is incorrectly configured.
- A slow disk fault has occurred.
- The Broker disk I/O is busy.
- Broker partitions are unevenly distributed, and hotspotting has occurred.

Handling Procedure

Check whether the number of threads used by the Broker instance to process requests is appropriate.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kafka**. On the page that is displayed, click **Configurations** and then **All Configurations**.

Step 2 Search for and check the value of **num.io.threads**. If the value is too small, increase it. You are advised to change the value to twice the number of CPU cores. The maximum value is **64**. Save the configuration.

Step 3 Click the **Instances** tab, select all Broker instances, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. Restart the instances during off-peak hours.

Step 4 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Check whether a slow disk fault has occurred.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the host name for which this alarm is generated.

Step 6 Check whether alarm **Slow Disk Fault** or **Disk Unavailable** is generated for the same node in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-12033 Slow Disk Fault** or **ALM-12063 Disk Unavailable**.
- If no, go to **Step 8**.

Step 7 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the Broker disk I/O is busy.

Step 8 Check whether alarm **Busy Broker Disk I/Os** exists on the node for which this alarm is generated in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-38009 Busy Broker Disk I/Os** and then go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether Broker partitions are evenly distributed and hotspotting has occurred.

Step 10 Choose **Cluster** > **Services** > **Kafka** > **Chart**, select **Partition** from the **Chart Category** area, zoom in **Number of Partitions-All Instances** in the upper right corner, and click **Distribution** to check whether partitions are evenly distributed on Broker.

- If yes, go to **Step 13**.
- If no, go to **Step 11**.

Step 11 Click the uneven distribution bar on the rightmost, and check whether the node obtained in **Step 5** is included in the unevenly distributed instances. If it is, perform data balancing.

Step 12 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.271 ALM-38014 Total Produce Request Latency Exceeds the Threshold

Alarm Description

The system checks the total latency of Produce requests on Broker instances every 30 seconds. This alarm is generated when the total latency of Produce requests on a Broker instance has exceeded the threshold for 10 consecutive times.

This alarm is cleared when the total latency of Produce requests is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38014	Critical (default threshold: 120000) Major (default threshold: 60000)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

The total latency of Produce requests on the Broker instance exceeds the threshold. For latency-sensitive services, a large number of service query requests may time out.

Possible Causes

- The number of threads used by the Broker instance to process requests is incorrectly configured.
- A slow disk fault has occurred.
- The Broker disk I/O is busy.
- Broker partitions are unevenly distributed, and hotspotting has occurred.

Handling Procedure

Check whether the number of threads used by the Broker instance to process requests is appropriate.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kafka**. On the page that is displayed, click **Configurations** and then **All Configurations**.

Step 2 Search for and check the value of **num.io.threads**. If the value is too small, increase it. You are advised to change the value to twice the number of CPU cores. The maximum value is **64**. Save the configuration.

Step 3 Click the **Instances** tab, select all Broker instances, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. Restart the instances during off-peak hours.

Step 4 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether a slow disk fault has occurred.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the host name for which this alarm is generated.

Step 6 Check whether alarm **Slow Disk Fault** or **Disk Unavailable** is generated for the same node in **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, rectify the fault by following the handling procedure of **ALM-12033 Slow Disk Fault** or **ALM-12063 Disk Unavailable**.
- If no, go to **Step 8**.

Step 7 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the Broker disk I/O is busy.

Step 8 Check whether alarm **Busy Broker Disk I/Os** exists on the node for which this alarm is generated in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-38009 Busy Broker Disk I/Os** and then go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether Broker partitions are evenly distributed and hotspotting has occurred.

Step 10 Choose **Cluster** > **Services** > **Kafka** > **Chart**, select **Partition** from the **Chart Category** area, zoom in **Number of Partitions-All Instances** in the upper right corner, and click **Distribution** to check whether partitions are evenly distributed on Broker.

- If yes, go to **Step 13**.
- If no, go to **Step 11**.

Step 11 Click the uneven distribution bar on the rightmost, and check whether the node obtained in **Step 5** is included in the unevenly distributed instances. If it is, perform data balancing.

Step 12 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M engineers and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.272 ALM-38015 Fetch Request Latency in the Request Queue Exceeds the Threshold

Alarm Description

The system checks the latency of Fetch requests on Broker instances in the request queue every 30 seconds. This alarm is generated when the latency of Fetch requests on a Broker instance in the request queue has exceeded the threshold for 10 consecutive times.

This alarm is cleared when the latency of Fetch requests in the request queue is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38015	Critical (default threshold: 30000) Major (default threshold: 10000)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The latency of Fetch requests on the Broker instance in the request queue exceeds the threshold. For latency-sensitive services, a large number of service query requests may time out.

Possible Causes

- The number of threads used by the Broker instance to process requests is incorrectly configured.
- A slow disk fault has occurred.
- The Broker disk I/O is busy.
- Broker partitions are unevenly distributed, and hotspotting has occurred.

Handling Procedure

Check whether the number of threads used by the Broker instance to process requests is appropriate.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kafka**. On the page that is displayed, click **Configurations** and then **All Configurations**.

Step 2 Search for and check the value of **num.io.threads**. If the value is too small, increase it. You are advised to change the value to twice the number of CPU cores. The maximum value is **64**. Save the configuration.

Step 3 Click the **Instances** tab, select all Broker instances, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. Restart the instances during off-peak hours.

Step 4 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether a slow disk fault has occurred.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the host name for which this alarm is generated.

Step 6 Check whether alarm **Slow Disk Fault** or **Disk Unavailable** is generated for the same node in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-12033 Slow Disk Fault** or **ALM-12063 Disk Unavailable**.
- If no, go to **Step 8**.

Step 7 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the Broker disk I/O is busy.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Check whether alarm **Busy Broker Disk I/Os** exists on the node for which this alarm is generated in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-38009 Busy Broker Disk I/Os** and then go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether Broker partitions are evenly distributed and hotspotting has occurred.

Step 10 Choose **Cluster** > **Services** > **Kafka** > **Chart**, select **Partition** from the **Chart Category** area, zoom in **Number of Partitions-All Instances** in the upper right corner, and click **Distribution** to check whether partitions are evenly distributed on Broker.

- If yes, go to **Step 13**.
- If no, go to **Step 11**.

Step 11 Click the uneven distribution bar on the rightmost, and check whether the node obtained in **Step 5** is included in the unevenly distributed instances. If it is, perform data balancing.

Step 12 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M engineers and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.273 ALM-38016 Total Fetch Request Latency Exceeds the Threshold

Alarm Description

The system checks the total latency of Fetch requests on Broker instances every 30 seconds. This alarm is generated when the total latency of Fetch requests on a Broker instance has exceeded the threshold for 10 consecutive times.

This alarm is cleared when the total latency of Fetch requests on the Broker instance is less than or equal to the threshold.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38016	Critical (default threshold: 60000) Major (default threshold: 30000)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

The total latency of Fetch requests on the Broker instance exceeds the threshold. For latency-sensitive services, a large number of service query requests may time out.

Possible Causes

- The number of threads used by the Broker instance to process requests is incorrectly configured.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- A slow disk fault has occurred.
- The Broker disk I/O is busy.
- Broker partitions are unevenly distributed, and hotspotting has occurred.

Handling Procedure

Check whether the number of threads used by the Broker instance to process requests is appropriate.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kafka**. On the page that is displayed, click **Configurations** and then **All Configurations**.

Step 2 Search for and check the value of **num.io.threads**. If the value is too small, increase it. You are advised to change the value to twice the number of CPU cores. The maximum value is **64**. Save the configuration.

Step 3 Click the **Instances** tab, select all Broker instances, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. Restart the instances during off-peak hours.

Step 4 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether a slow disk fault has occurred.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Location** field of the alarm details, view the host name for which this alarm is generated.

Step 6 Check whether alarm **Slow Disk Fault** or **Disk Unavailable** is generated for the same node in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-12033 Slow Disk Fault** or **ALM-12063 Disk Unavailable**.
- If no, go to **Step 8**.

Step 7 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the Broker disk I/O is busy.

Step 8 Check whether alarm **Busy Broker Disk I/Os** exists on the node for which this alarm is generated in **Step 5**.

- If yes, rectify the fault by following the handling procedure of **ALM-38009 Busy Broker Disk I/Os** and then go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 10**.

Check whether Broker partitions are evenly distributed and hotspotting has occurred.

Step 10 Choose **Cluster** > **Services** > **Kafka** > **Chart**, select **Partition** from the **Chart Category** area, zoom in **Number of Partitions-All Instances** in the upper right corner, and click **Distribution** to check whether partitions are evenly distributed on Broker.

- If yes, go to **Step 13**.
- If no, go to **Step 11**.

Step 11 Click the uneven distribution bar on the rightmost, and check whether the node obtained in **Step 5** is included in the unevenly distributed instances. If it is, perform data balancing.

Step 12 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M engineers and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.274 ALM-38017 Partition Reassignment Duration Exceeds the Threshold

Alarm Description

The system checks the partition reassignment time every 10 minutes. The check interval can be modified by the Kafka configuration **auto.reassign.check.interval.ms**. This alarm is generated when the partition reassignment triggered by a Broker scale-out takes more time than the threshold (1440 minutes by default, which can be modified by the Kafka configuration **reassignment.total.time.threshold**).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the partition workload balancing is complete.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38017	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the cluster service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

The Kafka service has unbalanced load for a long time, which may deteriorate the read and write performance.

Possible Causes

The amount of data in a partition to be migrated is too large, and the allowed triffic is low.

Handling Procedure

Step 1 Log in to the Kafka web UI.

1. Log in to FusionInsight Manager as a user who has the permission to access the Kafka web UI.
2. Choose **Cluster** > **Services** > **Kafka**.
3. On the right of **KafkaManager Web UI**, click the URL to access the Kafka web UI.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Click **Current Reassign Status** to view the partition reassignment tasks.

Step 3 Check the status of the partitions that are being migrated. If the progress remains unchanged for a long time, click **Modify Reassignment Throttle** to check whether the value of the **Throttle** parameter is too small.

- If yes, adjust the **Throttle** parameter during off-peak hours and click **OK** to accelerate the migration. Run the command in **Step 4**.
- If no, go to **Step 5**.

Step 4 Wait for 10 minutes and check whether the partition migration progress changes significantly.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.275 ALM-38018 Kafka Consumer Lag

Alarm Description

If you have configured a threshold to report Kafka consumer lag on the **Alarms** page of Kafka UI (there is no such rule by default), the system reports the alarm based on the following rules:

The system checks the topics subscribed to by all consumer groups every 60 seconds. This alarm is generated when the system detects that the difference (lag) between the consumption progress (offset) and the log end offset of the latest message generated in the partition is too large for five consecutive times, and the consumer log exceeds the threshold configured in the alarm rule.

This alarm is cleared when the system detects that the difference (lag) between the offsets is lower than the configured threshold for five consecutive times.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

This alarm applies only to MRS 3.5.0 or later.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
38018	User-defined by threshold (Critical/ Major)	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	ServiceName	Specifies the cluster service for which the alarm was generated.
	ConsumerGroup	Name of the Kafka consumer group for which the alarm is generated.
Additional Information	TopicName	Specifies the Kafka topic for which the alarm is generated.
	ConsumerLag	Specifies the number of messages yet to be consumed by the consumers in the Kafka topic for which the alarm is generated.

Impact on the System

Messages in Kafka topics are retained for a limited period (seven days by default). If messages are not consumed in time, data will be lost.

Possible Causes

- The new consumer group starts consuming messages from the beginning topic, leading to consumer lag.
- The threshold of the consumer lag alarm rule configured by the user is too small.
- The Kafka topic traffic increases sharply, and a large number of messages are generated in a short period of time.
- It takes a long time for the downstream system to process the Kafka messages in the topic.

Handling Procedure

Check whether the consumer group is new.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. View the alarm details. In the **Location** information area, view the name of the Kafka

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

consumer group for which the alarm is generated. In the **Additional Information** area, view the topic name.

Step 2 Check whether the consumer group is new.

- If yes, go to **Step 3**.

  NOTE

  In a new consumer group, the new consumer starts consuming messages from the beginning topic, which can cause a consumer lag alarm. This alarm is automatically cleared once the downstream consumer finishes processing the topic messages.

- If no, go to **Step 4**.

Step 3 Wait a moment and then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the alarm rule configuration is improper.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Kafka**. On the right of KafkaManager web UI, click the URL link to access the Kafka UI. Click **Alarms** and check whether the configured threshold of the consumer lag alarm is proper.

- If yes, go to **Step 6**.
- If no, reconfigure the threshold, save the configuration, and go to **Step 5**.

Step 5 Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the topic traffic increases sharply.

Step 6 On the Kafka UI, click **Topics** and check whether a large number of messages are generated in a short period of time.

- If yes, go to **Step 7**.

  NOTE

  If the alarm is caused by a soaring increase in topic traffic, the alarm is automatically cleared after the downstream system consumes topic messages.

- If no, go to **Step 8**.

Step 7 Wait a moment and then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether it takes a long time for the downstream system to process messages in the Kafka topic.

Step 8 Check whether the downstream system is consuming messages from the topic at a slow pace.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Analyze the reason why downstream jobs cannot quickly consume the topic messages and rectify the fault to accelerate the consumption. Wait 5 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **Kafka** for the target cluster.

Step 12 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.276 ALM-43001 Spark2x Service Unavailable

Alarm Description

The system checks the Spark2x service status every 300 seconds. This alarm is generated when the Spark2x service is unavailable.

This alarm is cleared when the Spark2x service recovers.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JobHistory2x is changed to JobHistory. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
43001	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The Spark tasks submitted by users fail to be executed.

Possible Causes

- The KrbServer service is abnormal.
- The LdapServer service is abnormal.
- ZooKeeper is abnormal.
- HDFS is abnormal.
- Yarn is abnormal.
- The corresponding Hive service is abnormal.
- The Spark2x assembly package is abnormal.
- The NameNode memory is insufficient.
- The memory of the Spark process is insufficient.

Handling Procedure

If the alarm is abnormal Spark2x assembly packet, the Spark packet is abnormal. Wait for about 10 minutes. The alarm is automatically cleared.

Check whether service unavailability alarms exist in services that Spark2x depends on.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**.

Step 2 Check whether the following alarms exist in the alarm list:

- ALM-25500 KrbServer Service Unavailable
- ALM-25000 LdapServer Service Unavailable
- ALM-13000 ZooKeeper Service Unavailable
- ALM-14000 HDFS Service Unavailable
- ALM-18000 Yarn Service Unavailable

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- ALM-16004 Hive Service Unavailable
- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Handle the alarms based on the troubleshooting methods provided in the alarm help.

After the alarm is cleared, wait a few minutes and check whether the alarm GuardianService Unavailable is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the NameNode memory is insufficient.

Step 4 Check whether the NameNode memory is insufficient.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Restart the NameNode to release the memory. Then, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Check whether the memory of the Spark process is insufficient.

Step 6 Check whether the memory of the Spark process is insufficient due to memory-related modifications.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Ensure that the memory of the Spark process is sufficient or expand the cluster capacity. Then, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 9 In the **Service** area, select the following nodes of the desired cluster. (Hive is the specific Hive service determined based on **ServiceName** in the alarm location information).

- KrbServer
- LdapServer

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- ZooKeeper
- HDFS
- Yarn
- Hive

Step 10 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.277 ALM-43006 Heap Memory Usage of the JobHistory2x Process Exceeds the Threshold

Description

The system checks the JobHistory2x Process status every 30 seconds. The alarm is generated when the heap memory usage of a JobHistory2x Process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JobHistory2x is changed to JobHistory. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43006	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the direct memory usage of the JobHistory2x process is too high, the performance deteriorates, and the process even becomes unavailable due to memory overflow. When it is unavailable, execution records of Spark tasks cannot be queried.

Possible Causes

The heap memory of the JobHistory2x Process is overused or the heap memory is inappropriately allocated.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43006**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JobHistory2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > **Memory > JobHistory2x Memory Usage Statistics** from the drop-down list box in the upper right corner and click **OK**. Check whether the used heap memory of the JobHistory2x Process reaches the threshold (default value is 95%) of the maximum heap memory specified for JobHistory2x.

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Spark2x** > **Instance**. Click **JobHistory2x** by which the alarm is reported to go to the **Dashboard** page, click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Memory > Statistics for the heap memory of the JobHistory2x Process**, and click **OK**. Based on the alarm generation time, check the values of the used heap memory of the JobHistory2x process in the corresponding period and obtain the maximum value.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On the FusionInsight Manager portal, choose **Cluster** > **Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JobHistory2x** > **Default**. The default value of **SPARK_DAEMON_MEMORY** is **4 GB**. You can change the value according to the following rules: Ratio of the maximum heap memory usage of the JobHistory2x to the **Threshold** of the **JobHistory2x Heap Memory Usage Statistics (JobHistory2x)** in the alarm period. If this alarm is generated occasionally after the parameter value is adjusted, increase the value by 0.5 times. If the alarm is frequently reported after the parameter value is adjusted, increase the value by 1 time.

NOTE

On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Thresholds** > **Spark2x** > **Memory** > **JobHistory2x Heap Memory Usage Statistics (JobHistory2x)** to view **Threshold**.

Step 5 Restart all JobHistory2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **Spark2x** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.278 ALM-43007 Non-Heap Memory Usage of the JobHistory2x Process Exceeds the Threshold

Description

The system checks the JobHistory2x Process status every 30 seconds. The alarm is generated when the non-heap memory usage of a JobHistory2x Process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JobHistory2x is changed to JobHistory. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43007	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the non-heap memory usage of the JobHistory2x process is too high, the performance deteriorates, and the process even becomes unavailable due to memory overflow. When it is unavailable, execution records of Spark tasks cannot be queried.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The non-heap memory of the JobHistory2x Process is overused or the non-heap memory is inappropriately allocated.

Procedure

Check non-heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43007**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JobHistory2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > **Memory** > **JobHistory2x Memory Usage Statistics** from the drop-down list box in the upper right corner and click **OK**, Check whether the used non-heap memory of the JobHistory2x Process reaches the threshold (default value is 95%) of the maximum non-heap memory specified for JobHistory2x.

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Spark2x** > **Instance**. Click **JobHistory2x** by which the alarm is reported to go to the **Dashboard** page, click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Memory > Statistics for the non-heap memory of the JobHistory2x Process**, and click **OK**. Based on the alarm generation time, check the values of the used non-heap memory of the JobHistory2x process in the corresponding period and obtain the maximum value.

Step 4 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JobHistory2x** > **Default**. You can change the value of **-XX:MaxMetaspaceSize** in **SPARK_DAEMON_JAVA_OPTS** according to the following rules: Ratio of the JobHistory2x non-heap memory usage to the **Threshold** of **JobHistory2x Non-Heap Memory Usage Statistics (JobHistory2x)** in the alarm period.

NOTE

On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Thresholds > Spark2x** > **Memory** >**JobHistory2x Non-Heap Memory Usage Statistics (JobHistory2x)** to view **Threshold**.

Step 5 Restart all JobHistory2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Spark2x** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.279 ALM-43008 The Direct Memory Usage of the JobHistory2x Process Exceeds the Threshold

Description

The system checks the JobHistory2x Process status every 30 seconds. The alarm is generated when the direct memory usage of a JobHistory2x Process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JobHistory2x is changed to JobHistory. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43008	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Name	Meaning
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the direct memory usage of the JobHistory2x process is too high, the performance deteriorates, and the process even becomes unavailable due to memory overflow. When it is unavailable, execution records of Spark tasks cannot be queried.

Possible Causes

The direct memory of the JobHistory2x Process is overused or the direct memory is inappropriately allocated.

Procedure

Check direct memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43008**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JobHistory2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > **Memory** > **JobHistory2x Memory Usage Statistics** from the drop-down list box in the upper right corner and click **OK**. Check whether the used direct memory of the JobHistory2x Process reaches the threshold (default value is 95%) of the maximum direct memory specified for JobHistory2x.

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Spark2x** > **Instance**. Click **JobHistory2x** by which the alarm is reported to go to the **Dashboard** page, click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Memory > Direct Memory of JobHistory2x**, and click **OK**. Based on the alarm generation time, check the values of the used direct memory of the JobHistory2x process in the corresponding period and obtain the maximum value.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On the FusionInsight Manager portal, choose **Cluster> Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JobHistory2x** > **Default**. The default value of **-XX:MaxDirectMemorySize** in **SPARK_DAEMON_JAVA_OPTS** is 512 MB. You can change the value according to the following rules: Ratio of the maximum direct memory usage of the JobHistory2x to the **Threshold** of the **JobHistory2x Direct Memory Usage Statistics (JobHistory2x)** in the alarm period. If this alarm is generated occasionally after the parameter value is adjusted, increase the value by 0.5 times. If the alarm is frequently reported after the parameter value is adjusted, increase the value by 1 time. It is recommended that the value be less than or equal to the value of **SPARK_DAEMON_MEMORY**.

NOTE

On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Thresholds > Spark2x** > **Memory** > **JobHistory2x Direct Memory Usage Statistics (JobHistory2x)** to view **Threshold**.

Step 5 Restart all JobHistory2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log > Download**.

Step 8 Select **Spark2x** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.280 ALM-43009 JobHistory2x Process GC Time Exceeds the Threshold

Description

The system checks the garbage collection (GC) time of the JobHistory2x Process every 60 seconds. This alarm is generated when the detected GC time exceeds the threshold (exceeds 12 seconds for three consecutive checks.) To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > **Spark2x** > **GC Time** > **Total GC time in milliseconds (JobHistory2x)**. This alarm is cleared when the JobHistory2x GC time is shorter than or equal to the threshold.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JobHistory2x is changed to JobHistory. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43009	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

The process performance deteriorates, and the process can even be unavailable. Historical execution records of Spark tasks cannot be queried.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The memory of JobHistory2x is overused, the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Procedure

Check the GC time.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43009**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JobHistory2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > **GC Time** > **Garbage Collection (GC) Time of JobHistory2x** from the drop-down list box in the upper right corner and click **OK** to check whether the GC time is longer than the threshold (default value: 12 seconds).

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JobHistory2x** > **Default**. The default value of **SPARK_DAEMON_MEMORY** is **4 GB**. You can change the value according to the following rules: If this alarm is generated occasionally, increase the value by 0.5 times. If the alarm is frequently reported, increase the value by 1 time.

Step 4 Restart all JobHistory2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 5 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager interface of active and standby clusters, choose **O&M** > **Log > Download**.

Step 7 Select **Spark2x** in the required cluster from the **Service**.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.281 ALM-43010 Heap Memory Usage of the JDBCServer2x Process Exceeds the Threshold

Description

The system checks the JDBCServer2x Process status every 30 seconds. The alarm is generated when the heap memory usage of a JDBCServer2x Process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JDBCServer2x is changed to JDBCServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43010	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the heap memory usage of the JDBCServer2x process is too high, the performance deteriorates, and even memory overflow occurs. As a result, the JDBCServer2x process is unavailable, and Spark JDBC tasks are slow or fail to run.

Possible Causes

The heap memory of the JDBCServer2x Process is overused or the heap memory is inappropriately allocated.

Procedure

Check heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43010**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JDBCServer2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > **Memory** > **JDBCServer2x Memory Usage Statistics** from the drop-down list box in the upper right corner and click **OK**. Check whether the used heap memory of the JDBCServer2x Process reaches the threshold (default value is 95%) of the maximum heap memory specified for JDBCServer2x.

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Spark2x** > **Instance**. Click **JDBCServer2x** by which the alarm is reported to go to the **Dashboard** page, click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Memory > Statistics for the heap memory of the JDBCServer2x Process**, and click **OK**. Based on the alarm generation time, check the values of the used heap memory of the JDBCServer2x process in the corresponding period and obtain the maximum value.

Step 4 On the FusionInsight Manager portal, choose **Cluster> Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JDBCServer2x** > **Tuning**. The default value of **SPARK_DRIVER_MEMORY** is 4 GB. You can change the value according to the following rules: Ratio of the maximum heap memory usage of the JDBCServer2x to the **Threshold** of the **JDBCServer2x Heap Memory Usage Statistics (JDBCServer2x)** in the alarm period. If this alarm is generated occasionally after the parameter value is adjusted, increase the value by 0.5 times. If the alarm is frequently reported after the parameter value is adjusted, increase the value by 1 time. In the case of large service volume and high concurrency, add instances.

NOTE

On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Thresholds > Spark2x** > **Memory** > **JDBCServer2x Heap Memory Usage Statistics (JDBCServer2x)** to view **Threshold**.

Step 5 Restart all JDBCServer2x instances.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Spark2x** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.282 ALM-43011 Non-Heap Memory Usage of the JDBCServer2x Process Exceeds the Threshold

Description

The system checks the JDBCServer2x Process status every 30 seconds. The alarm is generated when the non-heap memory usage of an JDBCServer2x Process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JDBCServer2x is changed to JDBCServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43011	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the non-heap memory usage of the JDBCServer2x process is too high, the performance deteriorates, and even memory overflow occurs. As a result, the JDBCServer2x process is unavailable, and Spark JDBC tasks are slow or fail to run.

Possible Causes

The non-heap memory of the JDBCServer2x Process is overused or the non-heap memory is inappropriately allocated.

Procedure

Check non-heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43011**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JDBCServer2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > **Memory** > **JDBCServer2x Memory Usage Statistics** from the drop-down list box in the upper right corner and click **OK**, Check whether the used non-heap memory of the JDBCServer2x Process reaches the threshold (default value is 95%) of the maximum non-heap memory specified for JDBCServer2x.

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Spark2x** > **Instance**. Click **JDBCServer2x** by which the alarm is reported to go to the **Dashboard** page, click the drop-down list in the upper right corner of the chart

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

area, choose **Customize** > **Memory > Statistics for the non-heap memory of the JDBCServer2x Process**, and click **OK**. Based on the alarm generation time, check the values of the used non-heap memory of the JDBCServer2x process in the corresponding period and obtain the maximum value.

Step 4 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JDBCServer2x** > **Tuning**. You can change the value of **-XX: MaxMetaspaceSize** in **spark.driver.extraJavaOptions** according to the following rules: Ratio of the JDBCServer2x non-heap memory usage to the **Threshold** of **JDBCServer2x Non-Heap Memory Usage Statistics (JDBCServer2x)** in the alarm period.

NOTE

On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Thresholds > Spark2x** > **Memory** > **JDBCServer2x Non-Heap Memory Usage Statistics (JDBCServer2x)** to view **Threshold**.

Step 5 Restart all JDBCServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log >Download**.

Step 8 Select **Spark2x** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.283 ALM-43012 Direct Heap Memory Usage of the JDBCServer2x Process Exceeds the Threshold

Description

The system checks the JDBCServer2x Process status every 30 seconds. The alarm is generated when the direct heap memory usage of a JDBCServer2x Process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JDBCServer2x is changed to JDBCServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43012	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Specifies the threshold triggering the alarm. If the current indicator value exceeds this threshold, the alarm is generated.

Impact on the System

If the direct memory usage of the JDBCServer2x process is too high, the performance deteriorates, and even memory overflow occurs. As a result, the JDBCServer2x process is unavailable, and Spark JDBC tasks are slow or fail to run.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The direct heap memory of the JDBCServer2x Process is overused or the direct heap memory is inappropriately allocated.

Procedure

Check direct heap memory usage.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43012**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Instance** and click the JDBCServer2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the Chart area and choose **Customize** > Memory > **JDBCServer2x Memory Usage Statistics** from the drop-down list box in the upper right corner and click **OK**. Check whether the used direct heap memory of the JDBCServer2x Process reaches the threshold of the maximum direct heap memory specified for JDBCServer2x.

- If yes, go to **Step 3**.
- If no, go to **Step 7**.

Step 3 On the FusionInsight Manager home page, choose **Cluster** > **Services** > **Spark2x** > **Instance**. Click **JDBCServer2x** by which the alarm is reported to go to the **Dashboard** page, click the drop-down list in the upper right corner of the chart area, choose **Customize** > **Memory > Direct Memory of JDBCServer2x**, and click **OK**. Based on the alarm generation time, check the values of the used direct memory of the JDBCServer2x process in the corresponding period and obtain the maximum value.

Step 4 On the FusionInsight Manager portal, choose **Cluster > Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JDBCServer2x** > **Tuning**. The default value of **-XX:MaxDirectMemorySize** in **spark.driver.extraJavaOptions** is 512 MB. You can change the value according to the following rules: Ratio of the maximum direct memory usage of the JDBCServer2x to the **Threshold** of the **JDBCServer2x Direct Memory Usage Statistics (JDBCServer2x)** in the alarm period. If this alarm is generated occasionally after the parameter value is adjusted, increase the value by 0.5 times. If the alarm is frequently reported after the parameter value is adjusted, increase the value by 1 time. It is recommended that the parameter value be less than or equal to the value of **SPARK_DRIVER_MEMORY**. In the case of large service volume and high service concurrency, you are advised to add instances.

NOTE

On the FusionInsight Manager home page, choose **O&M** > **Alarm** > **Thresholds > Spark2x** > **Memory** > **JDBCServer2x Direct Memory Usage Statistics (JDBCServer2x)** to view **Threshold**.

Step 5 Restart all JDBCServer2x instances.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On the FusionInsight Manager portal, choose **O&M** > **Log** > **Download**.

Step 8 Select **Spark2x** in the required cluster from the **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.284 ALM-43013 JDBCServer2x Process GC Time Exceeds the Threshold

Alarm Description

The system checks the garbage collection (GC) time of the JDBCServer2x Process every 60 seconds. This alarm is generated when the detected GC time exceeds the threshold (exceeds 12 seconds for three consecutive checks.) To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > **Spark2x** > **GC Time** > **Total GC time in milliseconds (JDBCServer2x)**. This alarm is cleared when the JDBCServer2x GC time is shorter than or equal to the threshold.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JDBCServer2x is changed to JDBCServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
43013	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service name for which the alarm is generated.
RoleName	Specifies the role name for which the alarm is generated.
HostName	Specifies the object (host ID) for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

If the GC duration exceeds the threshold, the performance of the JDBCServer2x process deteriorates, and the process can even be unavailable. As a result, Spark JDBC tasks are slow or fail to run.

Possible Causes

The memory of JDBCServer2x is overused, the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Handling Procedure

Check the GC time.

Step 1 On the FusionInsight Manager portal, choose **O&M > Alarm > Alarms** and select the alarm whose **ID** is **43013**. Check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Spark2x** > **Instances** and click the JDBCServer2x for which the alarm is generated to go to the **Dashboard** page. Click the drop-down menu in the **Chart** area, choose **Customize** > **GC Time** > **Garbage Collection (GC) Time of JDBCServer2x** from the drop-down list box in the upper right corner, and click **OK** to check whether the GC time is longer than the threshold (default value: 12 seconds).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 On the FusionInsight Manager portal, choose **Cluster** > **Services** > **Spark2x** > **Configurations**, and click **All Configurations**. Choose **JDBCServer2x** > **Default**. The default value of **SPARK_DRIVER_MEMORY** is 4 GB. If this alarm is generated occasionally, increase the value by 0.5 times. If the alarm is frequently reported, increase the value by 1 time. In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Restart all JDBCServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 5 After 10 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On the FusionInsight Manager interface of active and standby clusters, choose **O&M** > **Log** > **Download**.

Step 7 Select **Spark2x** in the required cluster from the **Service**.

Step 8 Click  in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

Related Information

None

9.11.285 ALM-43017 JDBCServer2x Process Full GC Number Exceeds the Threshold

Description

The system checks the number of Full garbage collection (GC) times of the JDBCServer2x process every 60 seconds. This alarm is generated when the detected Full GC number exceeds the threshold (exceeds 12 for three consecutive

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

checks.) You can change the threshold by choosing **O&M > Alarm** > **Thresholds** > **Spark2x** > **GC number** > **Full GC Number of JDBCServer2x**. This alarm is cleared when the Full GC number of the JDBCServer2x process is less than or equal to the threshold.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JDBCServer2x is changed to JDBCServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43017	Major	Yes

Parameters

Name	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the full GC times exceed the threshold, the performance of the JDBCServer2x process deteriorates, and the process can even be unavailable. As a result, Spark JDBC tasks become slow or even fail.

Possible Causes

The heap memory usage of the JDBCServer2x process is excessively large, or the heap memory is inappropriately allocated. As a result, Full GC occurs frequently.

Procedure

Check the number of Full GCs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm > Alarms**, select this alarm, and check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 Choose **Cluster** > **Services** > **Spark2x** > **Instance**. On the displayed page, click the JDBCServer2x for which the alarm is reported. On the **Dashboard** page that is displayed, click the drop-down menu in the Chart area and choose **Customize** > **GC Number** > **Full GC Number of JDBCServer2x** in the upper right corner and click **OK**. Check whether the number of Full GCs of the JDBCServer2x process is greater than the threshold (default value: 12).

- If it is, go to **Step 3**.
- If it is not, go to **Step 6**.

Step 3 Choose **Cluster** > **Services** > **Spark2x** > **Configurations** > **All Configurations**. On the displayed page, choose **JDBCServer2x** > **Tuning**. The default value of **SPARK_DRIVER_MEMORY** is **4 GB**. You can change the value according to the following rules: If this alarm is generated occasionally, increase the value by 0.5 times. If the alarm is frequently reported, increase the value by 1 time. In the case of large service volume and high concurrency, add instances.

Step 4 Restart all JDBCServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 5 After 10 minutes, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

Step 6 Log in to FusionInsight Manager, and choose **O&M** > **Log** > **Download**.

Step 7 Select **Spark2x** in the required cluster from the **Service** drop-down list.

Step 8 Click  in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.286 ALM-43018 JobHistory2x Process Full GC Number Exceeds the Threshold

Description

The system checks the number of Full garbage collection (GC) times of the JobHistory2x process every 60 seconds. This alarm is generated when the detected Full GC number exceeds the threshold (exceeds 12 for three consecutive checks.) You can change the threshold by choosing **O&M** > **Alarm > Thresholds** > **Spark2x** > **GC number** > **Full GC Number of JobHistory2x**. This alarm is cleared when the Full GC number of the JobHistory2x process is less than or equal to the threshold.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, JobHistory2x is changed to JobHistory. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Alarm Severity	Auto Clear
43018	Major	Yes

Parameters

Name	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The process performance deteriorates, and the process can even be unavailable. Historical execution records of Spark tasks cannot be queried.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The heap memory usage of the JobHistory2x process is excessively large, or the heap memory is inappropriately allocated. As a result, Full GC occurs frequently.

Procedure

Check the number of Full GCs.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm > Alarms**, select this alarm, and check the **RoleName** in **Location** and confirm the IP address of **HostName**.

Step 2 Choose **Cluster** > **Services** > **Spark2x** > **Instance**. On the displayed page, click the JobHistory2x for which the alarm is reported. On the **Dashboard** page that is displayed, click the drop-down menu in the Chart area and choose **Customize** > **GC Number** > **Full GC Number of JobHistory2x** in the upper right corner and click **OK**. Check whether the number of Full GCs of the JobHistory2x process is greater than the threshold (default value: 12).

- If it is, go to **Step 3**.
- If it is not, go to **Step 6**.

Step 3 Choose **Cluster** > **Services** > **Spark2x** > **Configurations** > **All Configurations**. On the displayed page, choose **JobHistory2x** > **Default.** The default value of **SPARK_DAEMON_MEMORY** is **4 GB**. You can change the value according to the following rules: If this alarm is generated occasionally, increase the value by 0.5 times. If the alarm is frequently reported, increase the value by 1 time.

Step 4 Restart all JobHistory2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 5 After 10 minutes, check whether the alarm is cleared.

- If it is, no further action is required.
- If it is not, go to **Step 6**.

Collect fault information.

Step 6 Log in to FusionInsight Manager, and choose **O&M** > **Log** > **Download**.

Step 7 Select **Spark2x** in the required cluster from the **Service**.

Step 8 Click  in the upper right corner. In the displayed dialog box, set **Start Date** and **End Date** to 10 minutes before and after the alarm generation time respectively and click **OK**. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

This alarm will be automatically cleared after the fault is rectified.

Related Information

None

9.11.287 ALM-43019 Heap Memory Usage of the IndexServer2x Process Exceeds the Threshold

Description

The system checks the IndexServer2x process status every 30 seconds. The alarm is generated when the heap memory usage of a IndexServer2x process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, IndexServer2x is changed to IndexServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Severity	Auto Clear
43019	Major	Yes

Parameters

Parameter	Description
Source	Identifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the heap memory usage of the IndexServer2x process is too high, the performance deteriorates, and even memory overflow occurs. As a result, the IndexServer2x process is unavailable, and Carbon tasks with indexing enabled are slow or fail to run.

Possible Causes

The heap memory of the IndexServer2x process is overused or the heap memory is inappropriately allocated.

Procedure

Check the heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose ID is **43019**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance**. Click the IndexServer2x that reported the alarm to go to the **Instance Status** page. Click the drop-down list in the upper right corner of the chart area, and choose **Customize** > **Memory** > **IndexServer2x Memory Usage Statistics** > **OK**. Check whether the heap memory used by the IndexServer2x process has reached the upper limit (95% by default).

- If the threshold is reached, go to **Step 3**.
- If the threshold is not reached, go to **Step 7**.

Step 3 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance**. Click the IndexServer2x that reported the alarm to go to the **Instance Status** page. Click the drop-down list in the upper right corner of the chart area, and choose **Customize** > **Memory** > **Statistics for the heap memory of the IndexServer2x process** > **OK**. Based on the alarm generation time, check the values of the used heap memory of the process in the corresponding period and obtain the maximum value.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Configurations** > **All Configuration** > **IndexServer2x** > **Tuning**. The default value of the **SPARK_DRIVER_MEMORY** parameter is **4 GB**. You can change the value based on the ratio of the maximum heap memory used by the IndexServer2x process to the threshold specified by **IndexServer2x Heap Memory Usage Statistics (IndexServer2x)** in the alarm period. If the alarm persists after the parameter value is changed, increase the value by 0.5 times. If the alarm is generated frequently, increase the value by one time.

NOTE

On FusionInsight Manager, you can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Spark2x** > **Memory** > **IndexServer2x Heap Memory Usage Statistics (IndexServer2x)** to view the threshold.

Step 5 Restart all IndexServer2x instances.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 Check whether the alarm is cleared in 10 minutes.

- If the alarm is cleared, no further action is required.
- If the alarm is not cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **Spark2x** for the target cluster.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Reference

None

9.11.288 ALM-43020 Non-Heap Memory Usage of the IndexServer2x Process Exceeds the Threshold

Description

The system checks the IndexServer2x process status every 30 seconds. The alarm is generated when the non-heap memory usage of the IndexServer2x process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, IndexServer2x is changed to IndexServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Severity	Auto Clear
43020	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
Source	Identifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the non-heap memory usage of the IndexServer2x process is too high, the performance deteriorates, and memory overflow even occurs. As a result, the IndexServer2x process is unavailable, and Carbon tasks with indexing enabled become slow or fail.

Possible Causes

The non-heap memory of the IndexServer2x process is overused or the non-heap memory is inappropriately allocated.

Procedure

Check non-heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose ID is **43020**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance**. Click the IndexServer2x that reported the alarm to go to the **Instance Status** page. Click the drop-down list in the upper right corner of the chart area, and choose **Customize** > **IndexServer2x Memory Usage Statistics** > **OK**. Check whether the non-heap memory used by the process has reached the upper limit (95% by default).

- If the threshold is reached, go to **Step 3**.
- If the threshold is not reached, go to **Step 7**.

Step 3 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance**. Click the IndexServer2x that reported the alarm to go to the **Dashboard** page. Click the drop-down list in the upper right corner of the chart area, and choose **Customize** > **Statistics for the non-heap memory of the IndexServer2x process** > **OK**. Based on the alarm generation time, check the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

values of the used non-heap memory of the process in the corresponding period and obtain the maximum value.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Configurations** > **All Configurations** > **IndexServer2x** > **Performance**. You can change the value of **-XX:MaxMetaspaceSize** in the **spark.driver.extraJavaOptions** parameter based on the ratio of the maximum non-heap memory used by the IndexServer2x process to the threshold specified by **IndexServer2x Non-Heap Memory Usage Statistics (IndexServer2x)** in the alarm period.

NOTE

On FusionInsight Manager, you can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Spark2x** > **Memory** > **IndexServer2x Non-Heap Memory Usage Statistics (IndexServer2x)** to view the threshold.

Step 5 Restart all IndexServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 Check whether the alarm is cleared in 10 minutes.

- If the alarm is cleared, no further action is required.
- If the alarm is not cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **Spark2x** for the target cluster.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.289 ALM-43021 Direct Memory Usage of the IndexServer2x Process Exceeds the Threshold

Description

The system checks the IndexServer2x process status every 30 seconds. The alarm is generated when the direct heap memory usage of a IndexServer2x process exceeds the threshold (95% of the maximum memory).

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, IndexServer2x is changed to IndexServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Severity	Auto Clear
43021	Major	Yes

Parameters

Parameter	Description
Source	Identifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the direct memory usage of the IndexServer2x process is too high, the performance deteriorates, and even memory overflow occurs. As a result, the IndexServer2x process is unavailable, and Carbon tasks with indexing enabled are slow or fail to run.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The direct heap memory of the IndexServer2x process is overused or the direct heap memory is inappropriately allocated.

Procedure

Check direct heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose ID is **43021**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance**. Click the IndexServer2x that reported the alarm to go to the **Dashboard** page. Click the drop-down list in the upper right corner of the chart area, and choose **Customize** > **IndexServer2x Memory Usage Statistics** > **OK**. Check whether the direct memory used by the process has reached the threshold.

- If the threshold is reached, go to **Step 3**.
- If the threshold is not reached, go to **Step 7**.

Step 3 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance**. Click the IndexServer2x that reported the alarm to go to the **Instance Status** page. Click the drop-down list in the upper right corner of the chart area, and choose **Customize** > **Memory** > **Direct Memory of IndexServer2x** > **OK**. Based on the alarm generation time, check the values of the used direct memory of the process in the corresponding period and obtain the maximum value.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Configurations** > **All Configurations** > **IndexServer2x** > **Tuning**. You can change the value of **XX:MaxDirectMemorySize** (the default value is 512 MB) in the **spark.driver.extraJavaOptions** parameter based on the ratio of the maximum direct memory used by the IndexServer2x process to the threshold specified by **IndexServer2x Direct Memory Usage Statistics (IndexServer2x)** in the alarm period. If the alarm persists after the parameter value is changed, increase the value by 0.5 times. If the alarm is generated frequently, increase the value by one time.

NOTE

On FusionInsight Manager, you can choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Spark2x** > **Memory** > **IndexServer2x Direct Memory Usage Statistics (IndexServer2x)** to view the threshold.

Step 5 Restart all IndexServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 6 Check whether the alarm is cleared in 10 minutes.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the alarm is cleared, no further action is required.
- If the alarm is not cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **Spark2x** for the target cluster.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Reference

None

9.11.290 ALM-43022 IndexServer2x Process GC Time Exceeds the Threshold

Description

The system checks the GC time of the IndexServer2x process every 60 seconds. This alarm is generated when the detected GC time exceeds the threshold (12 seconds) for three consecutive times. To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Spark2x** > **GC Time** > **IndexServer2x Total GC time**. This alarm is cleared when the IndexServer2x GC time is shorter than or equal to the threshold.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, IndexServer2x is changed to IndexServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

Attribute

Alarm ID	Severity	Auto Clear
43022	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
Source	Identifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the GC duration exceeds the threshold, the performance of the IndexServer2x process deteriorates, and the process can even be unavailable. As a result, Carbon tasks with indexing enabled are slow or fail to run.

Possible Causes

The heap memory of the IndexServer2x process is overused or the heap memory is inappropriately allocated. As a result, GC occurs frequently.

Procedure

Check the GC time.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose ID is **43022**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance** and click the IndexServer2x for which the alarm is generated to enter the **Dashboard** page. Click the drop-down menu in the **Chart** area and choose **Customize** > **GC Time** > **Garbage Collection (GC) Time of IndexServer2x** from the drop-down list box in the upper right corner and click **OK** to check whether the GC time is longer than the threshold (12 seconds by default).

- If the threshold is reached, go to **Step 3**.
- If the threshold is not reached, go to **Step 6**.

Step 3 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Configurations** > **All Configurations** > **IndexServer2x** > **Default**. The default value of the **SPARK_DRIVER_MEMORY** is 4 GB. You can increase the value 1.5 times to its default. If this alarm is still generated occasionally after the adjustment, increase the value by 0.5 times. Double the value if the alarm is reported frequently.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Restart all IndexServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 5 Check whether the alarm is cleared in 10 minutes.

- If the alarm is cleared, no further action is required.
- If the alarm is not cleared, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Spark2x** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Reference

None

9.11.291 ALM-43023 IndexServer2x Process Full GC Number Exceeds the Threshold

Description

The system checks the Full GC number of the IndexServer2x process every 60 seconds. This alarm is generated when the detected Full GC number exceeds the threshold (12) for three consecutive times. To change the threshold, choose **O&M** > **Alarm** > **Thresholds** > *Name of the desired cluster* > **Spark2x** > **GC Number** > **Full GC Number of IndexServer2x**. This alarm is cleared when the Full GC number of the IndexServer2x process is less than or equal to the threshold.

NOTE

In MRS 3.3.0-LTS and later versions, the Spark2x component is renamed Spark, and the role names in the component are also changed. For example, IndexServer2x is changed to IndexServer. Refer to the descriptions and operations related to the component name and role names in the document based on your MRS version.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Severity	Auto Clear
43023	Major	Yes

Parameters

Parameter	Description
Source	Identifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the GC times exceeds the threshold, the performance of the IndexServer2x process deteriorates, and the process can even be unavailable. As a result, Carbon tasks with indexing enabled are slow or fail to run.

Possible Causes

The heap memory of the IndexServer2x process is overused or the heap memory is inappropriately allocated. As a result, Full GC occurs frequently.

Procedure

Check the number of Full GCs.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose ID is **43023**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Instance** and click the IndexServer2x for which the alarm is generated to enter the **Dashboard** page. Click the drop-down menu in the chart area and choose **Customize** > **Full GC Number of IndexServer2x** from the drop-down list box in the upper right corner and click **OK** to check whether the full GC number is larger than the threshold (default value: 12).

- If the threshold is reached, go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the threshold is not reached, go to **Step 6**.

Step 3 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark2x** > **Configurations** > **All Configurations** > **IndexServer2x** > **Performance**. The default value of the **SPARK_DRIVER_MEMORY** is **4 GB**. You can change the value according to the following rules: If this alarm is generated occasionally, increase the value by 0.5 times. If the alarm is reported frequently, double the value. In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Restart all IndexServer2x instances.

NOTICE

When the instance is rebooted, it cannot be used and any tasks running on the current instance node will fail.

Step 5 Check whether the alarm is cleared in 10 minutes.

- If the alarm is cleared, no further action is required.
- If the alarm is not cleared, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Spark2x** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

After the fault is rectified, the system automatically clears this alarm.

Reference

None

9.11.292 ALM-43028 JDBCServer Session Overflow

Alarm Description

This alarm is generated when the JDBCServer process forwards requests and triffic control is triggered due to insufficient session resources. In this case, too many requests are sent to the JDBCServer process, exceeding the processing capability of the JDBCServer process.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
43028	Minor (default threshold: 9 for three consecutive times) Critical (default threshold: 12 for three consecutive times)	No

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

Impact on the System

Requests that cannot be processed are responded with a failure message.

Possible Causes

The JDBCServer process on the node is overloaded.

Handling Procedure

Check the source of the requests for connecting to the JDBCServer service.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and select the alarm whose ID is **43028**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **Spark** > **Instances**, click the JDBCServer for which this alarm is generated, and click **jdbcserver-audit** in the **Log** column on the left.

Step 3 Click the download button in the lower left corner to download the logs to the local PC.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Search for **UserIP** in the logs, collect statistics on the IP addresses of the clients that submit a large number of requests, and limit the traffic of these clients to reserve resources for other clients.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Spark** for the target cluster.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm needs to be cleared manually.

Related Information

None.

9.11.293 ALM-43029 JDBCServer Job Submission Timed Out

Alarm Description

After a user submits a JDBC job, the system attempts to create a JDBCServer process and establish a session connection. This alarm is generated if the preset thresholds are exceeded before the connection is established. There are two configuration parameters affecting alarm triggering:

- **spark.thriftserver.proxy.create.session.monitor.enabled**: whether to enable the alarm function. The default value is **true** for the cluster.
- **spark.thriftserver.proxy.create.session.timeout.threshold**: Maximum time allowed for submitting a JDBC job. This alarm is reported when the system detects that the job does not start after the threshold is exceeded. The unit is second. The default value is 180s.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
43029	Major	No

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	User_Queue	Name of the user who submits the alarm and the queue for which the alarm is generated.

Impact on the System

The JDBC job submission time increases due to high system load, which may also affect the job execution efficiency. The job can start properly after this alarm is reported because the detection is asynchronous.

Possible Causes

The JDBCServer on the node is overloaded. The cluster health reflected by the system metrics and job execution status is not good.

Handling Procedure

Check the JDBCServer instance for which the alarm is generated.

Step 1 On FusionInsight Manager, choose O&M > Alarm > Alarms and select the alarm whose ID is 43029. Check the role name and the IP address of the host where the alarm is generated in **Location**. Check the username and queue in **Additional Information**.

Re-executing Affected JDBCServer Jobs

Step 2 Choose **Cluster** > **Services** > **Yarn** > **ResourceManager (Active)** to log in to the YARN web UI. Find the corresponding application based on the username and queue name in **Additional Information** and check whether the job submission is affected based on the driver log and Spark UI. Confirm and record the affected job to execute it again.

Step 3 On FusionInsight Manager, choose **Cluster** > **Services** > **Spark** > **Instances**, click the JDBCServer for which this alarm is generated, and choose **More** > **Restart Instance**.

Step 4 Choose **O&M** > **Alarm** > **Alarms**, search for the reported alarm, and click **Clear** in the **Operation** column.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Execute the affected job and check whether the alarm is triggered again for the job.

- If no, no further action is required.
- If yes, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Spark** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm needs to be cleared manually.

Related Information

None.

9.11.294 ALM-44004 Presto Coordinator Resource Group Queuing Tasks Exceed the Threshold

Alarm Description

This alarm is generated when the system detects that the number of queuing tasks in a resource group exceeds the threshold. The system queries the number of queuing tasks in a resource group through the JMX interface. You can choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Presto** > **resource-groups** to configure a resource group. You can choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Coordinator** > **Customize** > **resourceGroupAlarm** to configure the threshold of each resource group.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
44004	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

If the number of queuing tasks in a resource group exceeds the threshold, a large number of tasks may be in the queuing state. The Presto task time exceeds the expected value. When the number of queuing tasks in a resource group exceeds the maximum number (**maxQueued**) of queuing tasks in the resource group, new tasks cannot be executed.

Possible Causes

The resource group configuration is improper or too many tasks in the resource group are submitted.

Handling Procedure

Step 1 Choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Presto** > **resource-groups** to adjust the resource group configuration.

Step 2 You can choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Coordinator** > **Customize** > **resourceGroupAlarm** to modify the threshold of each resource group.

Step 3 Collect the fault information.

1. Log in to the cluster node based on the host name in the fault information and query the number of queuing tasks based on **Resource Group** in the additional information on the Presto client.
2. Log in to the cluster node based on the host name in the fault information, view the **/var/log/Bigdata/nodeagent/monitorlog/monitor.log** file, and search for resource group information to view the monitoring collection information of the resource group.
3. Contact O&M personnel and send the collected logs.

----End

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.295 ALM-44005 Presto Coordinator Process GC Time Exceeds the Threshold

Description

The system collects GC time of the Presto Coordinator process every 30 seconds. This alarm is generated when the GC time exceeds the threshold (exceeds 5 seconds for three consecutive times). You can change the threshold by choosing **System** > **Configure Alarm Threshold** > **Service** > **Presto** > **Coordinator** > **Presto Process Garbage Collection Time** > **Garbage Collection Time of the Coordinator Process** on MRS Manager. This alarm is cleared when the Coordinator process GC time is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
44005	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

If the GC time of the Coordinator process is too long, the Coordinator process running performance will be affected and the Coordinator process will even be unavailable.

Possible Causes

The heap memory of the Coordinator process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Procedure

Step 1 Check the GC time.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **44005** and then check the role name in **Location** and confirm the IP address of the instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Choose **Components** > **Presto** > **Instances** > **Coordinator** (business IP address of the instance for which the alarm is generated) > **Customize** > **Presto Garbage Collection Time**. Click **OK** to view the GC time.
4. Check whether the GC time of the Coordinator process is longer than 5 seconds.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Presto** > **Service Configuration**, and switch **Basic** to **All**. Choose **Presto** > **Coordinator**. Increase the value of **-Xmx** (maximum heap memory) in the **JAVA_OPTS** parameter based on the site requirements.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On FusionInsight Manager, choose **O&M** > **Log** > **Download**.
2. Expand the **Service** drop-down list, select **Presto** for the target cluster, and click **OK**.
3. Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.
4. Contact the O&M personnel and send the collected logs.

----End

Reference

None

9.11.296 ALM-44006 Presto Worker Process GC Time Exceeds the Threshold

Description

The system collects GC time of the Presto Worker process every 30 seconds. This alarm is generated when the GC time exceeds the threshold (exceeds 5 seconds for three consecutive times). You can change the threshold by choosing **System** > **Configure Alarm Threshold** > **Service** > **Presto** > **Worker** > **Presto Garbage Collection Time** > **Garbage Collection Time of the Worker Process** on MRS Manager. This alarm is cleared when the Worker process GC time is shorter than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
44006	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

If the GC time of the Worker process is too long, the Worker process running performance will be affected and the Worker process will even be unavailable.

Possible Causes

The heap memory of the Worker process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Procedure

Step 1 Check the GC time.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **44006**. Then check the role name in **Location** and confirm the IP address of the instance.
3. Choose **Components** > **Presto** > **Instances** > **Worker** (business IP address of the instance for which the alarm is generated) > **Customize** > **Presto Garbage Collection Time**. Click **OK** to view the GC time.
4. Check whether the GC time of the Worker process is longer than 5 seconds.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Presto** > **Service Configuration**, and switch **Basic** to **All**, and choose **Presto** > **Worker**. Increase the value of **-Xmx** (maximum heap memory) in the **JAVA_OPTS** parameter based on the site requirements.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Expand the **Service** drop-down list, select **Presto** for the target cluster, and click **OK**.
3. Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Contact the O&M personnel and send the collected logs.

----End

Reference

None

9.11.297 ALM-45000 HetuEngine Service Unavailable

Alarm Description

The system checks the HetuEngine service status every 300 seconds. This alarm is generated when the HetuEngine service is unavailable.

This alarm is cleared when the HetuEngine service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45000	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

FusionInsight Manager cannot be used to perform operations on the HetuEngine cluster, and HetuEngine functions are unavailable.

Possible Causes

- The KrbServer service is abnormal.
- The ZooKeeper service is abnormal.
- The HDFS service is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The Yarn service is abnormal.
- The DBService service is abnormal.
- The Hive service is abnormal.
- There are no HSBroker instances in HetuEngine.

Handling Procedure

Check the KrbServer service status.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarm**.

Step 2 In the alarm list, check whether the "ALM-25500 KrbServer Service Unavailable" alarm is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Clear the alarm according to the alarm help in **ALM-25500 KrbServer Service Unavailable**.

Step 4 In the alarm list, check whether the alarm "ALM-45000 HetuEngine Service Unavailable" is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the ZooKeeper service status.

Step 5 In the alarm list, check whether the alarm "ALM-12007 Process Fault" is generated.

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 In the alarm list, click ⌄ in the row that contains the "Process Fault" alarm. Check whether the name of the service for which the alarm is generated is ZooKeeper in **Location Information**.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Clear the alarm according to the alarm help in **ALM-12007 Process Fault**.

Step 8 In the alarm list, check whether the alarm "ALM-45000 HetuEngine Service Unavailable" is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check the HDFS service status.

Step 9 In the alarm list, check whether the "ALM-14000 HDFS Service Unavailable" alarm is generated.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Clear the alarm according to the alarm help in **ALM-14000 HDFS Service Unavailable**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 11 In the alarm list, check whether the "ALM-45000 HetuEngine Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Check the YARN service status.

Step 12 In the alarm list, check whether the "ALM-18000 YARN Service Unavailable" alarm is generated.

- If yes, go to **Step 13**.
- If no, go to **Step 15**.

Step 13 Clear the alarm according to the alarm help in **ALM-18000 YARN Service Unavailable**.

Step 14 In the alarm list, check whether the "ALM-45000 HetuEngine Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Check the DBService service status.

Step 15 In the alarm list, check whether the "ALM-27001 DBService Service Unavailable" alarm is generated.

- If yes, go to **Step 16**.
- If no, go to **20**.

Step 16 Clear the alarm according to the alarm help in **ALM-27001 DBService Service Unavailable**.

Step 17 In the alarm list, check whether the "ALM-45000 HetuEngine Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **20**.

Check the Hive service status.

Step 18 In the alarm list, check whether the "ALM-16004 Hive Service Unavailable" alarm is generated.

- If yes, go to **Step 19**.
- If no, go to **20**.

Step 19 Clear the alarm according to the alarm help in **ALM-16004 Hive Service Unavailable**.

Step 20 In the alarm list, check whether the "ALM-45000 HetuEngine Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **20**.

Check whether there are no HSBroker instances in HetuEngine.

Step 21 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HetuEngine**. On the page that is displayed, click the **Instance** tab.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 22 Check whether there are no HSBroker instances.

- If yes, click **Add Instance** to add one.
- If no, go to **23**.

Step 23 In the alarm list, check whether the "ALM-45000 HetuEngine Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **23**.

Check the network connection between HetuEngine and ZooKeeper, HDFS, YARN, DBService, and Hive.

Step 24 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **HetuEngine**. On the page that is displayed, click the **Instance** tab.

Step 25 Click the host name in the **HSBroker** row and record the management IP address in the **Basic Information** area.

Step 26 Log in to the host where HSBroker resides as user **omm** using the IP address obtained in **Step 25**.

Step 27 Run the **ping** command to check whether the network connection between the host where HSBroker resides and the hosts where ZooKeeper, HDFS, Yarn, DBService, and Hive reside is in the normal state.

- If yes, go to **Step 30**.
- If no, go to **Step 28**.

Step 28 Contact the network administrator to restore the network.

Step 29 In the alarm list, check whether the "ALM-45000 HetuEngine Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 30**.

Collect fault information.

Step 30 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 31 Expand the **Service** drop-down list. In the **Services** dialog box that is displayed, select **HetuEngine** under the target cluster name, and click **OK**.

Step 32 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 33 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 34 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the fault is rectified, the system automatically clears this alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.298 ALM-45001 Faulty HetuEngine Compute Instances

Alarm Description

The system checks the HetuEngine compute instance status every 60 seconds. This alarm is generated when a HetuEngine compute instance is faulty.

This alarm is cleared when all faulty HetuEngine compute instances are restored.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45001	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

SQL tasks submitted to the faulty compute instance of HetuEngine will fail.

Possible Causes

- The HDFS service is abnormal.
- The Yarn service is abnormal.
- Yarn queue resources are insufficient.
- The process of compute instances is faulty.
- The number of Worker progresses in the faulty compute instance is 0.

Handling Procedure

Check the HDFS service status.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 In the alarm list, check whether the "ALM-14000 HDFS Service Unavailable" alarm is generated.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Clear the alarm according to the alarm help in **ALM-14000 HDFS Service Unavailable**.

Step 3 In the alarm list, check whether the "ALM-45001 Faulty HetuEngine Compute Instances" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check the YARN service status.

Step 4 In the alarm list, check whether the "ALM-18000 YARN Service Unavailable" alarm is generated.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Clear the alarm according to the alarm help in **ALM-18000 YARN Service Unavailable**.

Step 6 In the alarm list, check whether the "ALM-45001 Faulty HetuEngine Compute Instances" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check the YARN queue resource status.

Step 7 In the alarm list, check whether the "ALM-18022 Insufficient YARN Queue Resources" alarm is generated.

- If yes, go to **8**.
- If no, go to **Step 10**.

Step 8 Clear the alarm according to the alarm help in **ALM-18022 Insufficient YARN Queue Resources**.

Step 9 In the alarm list, check whether the "ALM-45001 Faulty HetuEngine Compute Instances" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check the HetuEngine compute instance status.

Step 10 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI and choose **Cluster** > **Services > HetuEngine**.

Step 11 In the **Basic Information** area on the **Dashboard** tab page, click the link next to **HSConsole WebUI** to access the HSConsole page.

Step 12 On the compute instance page, check whether any compute instances are in the **FAULT** state.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 13 Check whether there are compute instances with no Worker progress.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 In the **Operation** column of the target compute instance, click **Start** and wait until the instance is started.

Step 15 In the alarm list, check whether the "ALM-45001 Faulty HetuEngine Compute Instances" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Collect fault information.

Step 16 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 17 Expand the **Service** drop-down list. In the **Services** dialog box that is displayed, select **HetuEngine** under the target cluster name, and click **OK**.

Step 18 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 19 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 20 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.299 ALM-45004 Tasks Stacked on HetuEngine Compute Instance

Alarm Description

The system checks the number of running tasks on a HetuEngine compute instance every 30 seconds. This alarm is generated when the number of running tasks is greater than 50.

This alarm is cleared when the number of tasks running on the HetuEngine compute instance is no more than 50.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45004	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Running Queries Backlog	Specifies the tenant name of the compute instance for which the alarm is generated and how much the threshold is exceeded.

Impact on the System

The performance of the compute instance deteriorates and the SQL response becomes slow.

Possible Causes

- The compute instance specification is too small.
- Large SQL tasks occupy too many compute resources. No resource is available for other tasks, and the compute instance cannot respond quickly. As a result, tasks are stacked.

Handling Procedure

Check whether compute instance resources are properly configured.

Step 1 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **Tasks Stacked on HetuEngine Compute Instance**, check the **Additional Information** of the alarm, and view and record the tenant name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Choose **Cluster** > **Services** > HetuEngine. In the **Basic Information** area in the **Dashboard** tab, click the link next to **HSConsole Web UI**. The HSConsole page is displayed.

Step 4 On the **Compute Instance** page, click **Configure** in the **Operation** column of the tenant to which the compute instance belongs. Check whether the resource configured for the compute instance is proper. (The minimum resources are used by default. You can adjust the configuration based on the site requirements.)

- If yes, go to **Step 8**.
- If no, go to **Step 5**.

Step 5 Return to the compute instance list, click **Stop Instances** in the **Operation** column, and stop instances as prompted.

NOTICE

Tasks submitted to the stopped compute instances will be interrupted.

Step 6 Click **Configure**, add resources to the target compute instance based on the site requirements, and click **OK**. Click **Start Instances** and start instances as prompted.

Step 7 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether there are large SQL tasks.

Step 8 On the **Compute Instances** page, expand the instances of the tenant and click **LINK** in the **WebUI** column of a compute instance to view the status of all tasks.

Step 9 In the **Sort** column, select **Execution Time** to sort the running tasks and check whether there are tasks that have been running for hours.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 End the tasks that have been running for a long time based on service requirement and optimize the service SQL statements.

Step 11 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 13 Expand the **Service** drop-down list, select **HetuEngine** for the target cluster, and click **OK**.

Step 14 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.300 ALM-45005 CPU Usage of HetuEngine Compute Instance Exceeded the Threshold

Alarm Description

The system checks the average CPU usage of HetuEngine compute instances every 30 seconds. This alarm is generated when the average CPU usage of the instances is greater than 90%.

This alarm is cleared when the CPU usage of the HetuEngine compute instances is no more than 90%.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45005	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Cpu Usage Exceeds The Threshold	Specifies the tenant name of the compute instance for which the alarm is generated and how much the threshold is exceeded.

Impact on the System

The performance of the compute instances deteriorates and the response to SQL statements becomes slow.

Possible Causes

- The compute instance specification is too small.
- Large SQL tasks occupy too many compute resources. No resource is available for other tasks, and the compute instance cannot respond quickly. As a result, tasks are stacked.

Handling Procedure

Check whether compute instance resources are properly configured.

Step 1 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **Tasks Stacked on HetuEngine Compute Instance**, check the **Additional Information** of the alarm, and view and record the tenant name for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > HetuEngine. In the **Basic Information** area in the **Dashboard** tab, click the link next to **HSConsole Web UI**. The HSConsole page is displayed.

Step 4 On the **Compute Instance** page, click **Configure** in the **Operation** column of the tenant to which the compute instance belongs. Check whether the resource configured for the compute instance is proper. (The minimum resources are used by default. You can adjust the configuration based on the site requirements.)

- If yes, go to **Step 8**.
- If no, go to **Step 5**.

Step 5 Return to the compute instance list, click **Stop Instances** in the **Operation** column, and stop instances as prompted.

NOTICE

Tasks submitted to the stopped compute instances will be interrupted.

Step 6 Click **Configure**, add resources to the target compute instance based on the site requirements, and click **OK**. Click **Start Instances** and start instances as prompted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether there are large SQL tasks.

Step 8 On the **Compute Instances** page, expand the instances of the tenant and click **LINK** in the **WebUI** column of a compute instance to view the status of all tasks.

Step 9 In the **Sort** column, select **Execution Time** to sort the running tasks and check whether there are tasks that have been running for hours.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 End the tasks that have been running for a long time based on service requirement and optimize the service SQL statements.

Step 11 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 13 Expand the **Service** drop-down list, select **HetuEngine** for the target cluster, and click **OK**.

Step 14 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M engineers and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.301 ALM-45006 Memory Usage of a HetuEngine Compute Instance Exceeded the Threshold

Alarm Description

The system checks the memory usage of HetuEngine compute instances every 30 seconds. This alarm is generated when the memory usage of the instance is greater than 80%.

This alarm is cleared when the memory usage of the HetuEngine compute instance is no more than 80%.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45006	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Memory Usage Exceeds The Threshold	Specifies the tenant name of the compute instance for which the alarm is generated and how much the threshold is exceeded.

Impact on the System

The performance of the compute instance deteriorates and the response to service SQL statements becomes slow.

Possible Causes

- The compute instance specification is too small.
- Large SQL tasks occupy too many compute resources. No resource is available for other tasks, and the compute instance cannot respond quickly. As a result, tasks are stacked.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check whether compute instance resources are properly configured.

Step 1 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **Tasks Stacked on HetuEngine Compute Instance**, check the **Additional Information** of the alarm, and view and record the tenant name for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > HetuEngine. In the **Basic Information** area in the **Dashboard** tab, click the link next to **HSConsole Web UI**. The HSConsole page is displayed.

Step 4 On the **Compute Instance** page, click **Configure** in the **Operation** column of the tenant to which the compute instance belongs. Check whether the resource configured for the compute instance is proper. (The minimum resources are used by default. You can adjust the configuration based on the site requirements.)

- If yes, go to **Step 8**.
- If no, go to **Step 5**.

Step 5 Return to the compute instance list, click **Stop Instances** in the **Operation** column, and stop instances as prompted.

NOTICE

Tasks submitted to the stopped compute instances will be interrupted.

Step 6 Click **Configure**, add resources to the target compute instance based on the site requirements, and click **OK**. Click **Start Instances** and start instances as prompted.

Step 7 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether there are large SQL tasks.

Step 8 On the **Compute Instances** page, expand the instances of the tenant and click **LINK** in the **WebUI** column of a compute instance to view the status of all tasks.

Step 9 In the **Sort** column, select **Execution Time** to sort the running tasks and check whether there are tasks that have been running for hours.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 End the tasks that have been running for a long time based on service requirement and optimize the service SQL statements.

Step 11 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 13 Expand the **Service** drop-down list, select **HetuEngine** for the target cluster, and click **OK**.

Step 14 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M engineers and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.302 ALM-45007 Number of Workers of a HetuEngine Compute Instance Is Less Than the Threshold

Alarm Description

The system checks the number of Workers of a HetuEngine compute instance every 60 seconds. This alarm is generated when the number of Workers is less than 80% of the initial value.

This alarm is cleared when the number of Workers running on the HetuEngine compute instance is no less than 80% of the initial value.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45007	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Worker Less Threshold	Specifies the tenant name of the compute instance for which the alarm is generated and how much the threshold is exceeded.

Impact on the System

The performance of the compute instance deteriorates and the SQL response becomes slow.

Possible Causes

- YARN queue resources are insufficient.
- A large number of tasks are running, causing OMM memory overflow on Worker nodes. As a result, the number of Worker nodes decreases.

Handling Procedure

Check whether YARN resource queue resources are sufficient.

Step 1 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **Number of Workers of a HetuEngine Compute Instance Is Less Than the Threshold**, check the **Additional Information** of the alarm, and view and record the tenant name for which the alarm is generated.

Step 3 Click **Tenant Resources**, select the tenant of the compute instance, and check whether the resource quota of the tenant is sufficient.

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Increase the maximum percentage of the tenant's resources based on the actual usage.

Step 5 Wait 5 to 10 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether there is a large number of running tasks.

Step 6 Choose **Cluster** > **Services** > **HetuEngine**.

Step 7 In the **Basic Information** area on the **Dashboard** tab page, click the link next to **HSConsole Web UI** to access the HSConsole page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 On the **Compute Instances** page, expand the instances of the tenant and click **LINK** in the **WebUI** column of a compute instance to view the status of all tasks.

Step 9 Check whether the number of running tasks exceeds 50.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Reduce the number of jobs submitted at a time or add compute instance resources based on service requirements.

Step 11 Wait 5 to 10 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 13 Expand the **Service** drop-down list, select **HetuEngine** for the target cluster, and click **OK**.

Step 14 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.303 ALM-45008 Query Latency of HetuEngine Compute Instances Exceeds the Threshold

Alarm Description

The system checks the query latency of HetuEngine compute instances every 30 seconds. This alarm is generated when the query latency of a HetuEngine compute instance is greater than or equal to 60 seconds.

This alarm is cleared when the query latency is less than 60s.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45008	Warning	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Running Queries Delay	Specifies the tenant name of the compute instance for which the alarm was generated and how much the threshold is exceeded.

Impact on the System

If the query latency of the HetuEngine compute instance exceeds the threshold, the SQL response of the service is slow.

Possible Causes

- The compute instance specification is too small.
- Large SQL tasks occupy too many compute resources. No resource is available for other tasks, and the compute instance cannot respond quickly. As a result, tasks are stacked.

Handling Procedure

Check whether compute instance resources are properly configured.

Step 1 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI.

Step 2 Choose **O&M** > **Alarm** > **Alarms**. In the right pane, click alarm **Query Latency of HetuEngine Compute Instances Exceeds the Threshold**, and view and record the tenant name in **Additional Information**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Choose **Cluster** > **Services** > HetuEngine. In the **Basic Information** area on the **Dashboard** page, click the link next to **HSConsole Web UI**. The HSConsole page is displayed.

Step 4 On the **Compute Instance** page, click **Configure** in the **Operation** column of the tenant to which the compute instance belongs. Check whether the resources configured for the compute instance are proper. (The minimum resources are used by default. You can adjust the configuration based on site requirements.)

- If yes, go to **Step 8**.
- If no, go to **Step 5**.

Step 5 Return to the compute instance list, click **Stop Instances** in the **Operation** column, and stop instances as prompted.

NOTICE

Tasks submitted to the stopped compute instances will be interrupted.

Step 6 Click **Configure**, add resources to the target compute instance based on site requirements, and click **OK**. Click **Start Instances** and start instances as prompted.

Step 7 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether there are large SQL tasks.

Step 8 On the **Compute Instances** page, expand the instances of the tenant and click **LINK** in the **WebUI** column of a compute instance to view the status of all tasks.

Step 9 In the **Sort** column, select **Execution Time** to sort the running tasks and check whether there are tasks that have been running for hours.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Stop the tasks that have been running for a long time based on service requirement and optimize the service SQL statements.

Step 11 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 13 Expand the **Service** drop-down list, select **HetuEngine** for the target cluster, and click **OK**.

Step 14 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 15 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.304 ALM-45009 Task Failure Rate of HetuEngine Compute Instances Exceeds the Threshold

Alarm Description

The system checks the task failure rate of HetuEngine compute instances every 30 seconds. This alarm is generated when the task failure rate of a HetuEngine compute instance is greater than or equal to 50%.

This alarm is cleared when the task failure rate is less than 50%.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45009	Warning	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Task Failure Rate	Specifies the tenant name of the compute instance for which the alarm was generated and how much the threshold is exceeded.

Impact on the System

If the task failure rate of HetuEngine compute instances is too high, service running is adversely affected. You need to locate and rectify the fault in a timely manner.

Possible Causes

- The compute instance specification is too small.
- Large SQL tasks occupy too many compute resources. No resource is available for other tasks, and the compute instance cannot respond quickly. As a result, tasks are stacked.
- The possible causes of SQL execution failures include incorrect syntax, incorrect permissions, and improper parameter settings.

Handling Procedure

Check whether compute instance resources are properly configured.

Step 1 Log in to FusionInsight Manager as an administrator who can access the HetuEngine web UI.

Step 2 Choose **O&M** > **Alarm** > **Alarms**. In the right pane, click alarm **Task Failure Rate of HetuEngine Compute Instances Exceeds the Threshold**, and view and record the tenant name in **Additional Information**.

Step 3 Choose **Cluster** > **Services** > HetuEngine. In the **Basic Information** area on the **Dashboard** page, click the link next to **HSConsole Web UI**. The HSConsole page is displayed.

Step 4 On the **Compute Instance** page, click **Configure** in the **Operation** column of the tenant to which the compute instance belongs. Check whether the resources configured for the compute instance are proper. (The minimum resources are used by default. You can adjust the configuration based on site requirements.)

- If yes, go to **Step 8**.
- If no, go to **Step 5**.

Step 5 Return to the compute instance list, click **Stop Instances** in the **Operation** column, and stop instances as prompted.

NOTICE

Tasks submitted to the stopped compute instances will be interrupted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Click **Configure**, add resources to the target compute instance based on site requirements, and click **OK**. Click **Start Instances** and start instances as prompted.

Step 7 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether there are large SQL tasks.

Step 8 On the **Compute Instances** page, expand the instances of the tenant and click **LINK** in the **WebUI** column of a compute instance to view the status of all tasks.

Step 9 In the **Sort** column, select **Execution Time** to sort the running tasks and check whether there are tasks that have been running for hours.

- If yes, go to **Step 10**.
- If no, select all failed tasks in the **Failed** column, collect detailed error information about SQL running in the computing instance, and go to **Step 14**.

Step 10 Stop the tasks that have been running for a long time based on service requirement and optimize the service SQL statements.

Step 11 Wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, select all failed tasks in the **Failed** column, collect detailed error information about SQL running in the computing instance, and go to **Step 14**.

Locate the stack where the SQL statement fails to be executed.

Step 12 Choose **Compute Instance**. In the right pane, click **LINK** in the **WebUI** column of the compute instance. On the details page, click **Failed** and select all error types to filter out all failed SQL statements.

Step 13 Click **Query ID** on the left of the SQL statement that reports the error. On the details page, view the error stack information and go to **Step 14**.

Collect fault information.

Step 14 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 15 Expand the **Service** drop-down list, select **HetuEngine** for the target cluster, and click **OK**.

Step 16 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 17 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 18 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.305 ALM-45175 Average Time for Calling OBS Metadata APIs Is Greater than the Threshold

Alarm Description

The system checks whether the average duration for calling OBS metadata APIs is greater than the threshold every 30 seconds. This alarm is generated when the number of consecutive times that the average time exceeds the specified threshold is greater than the number of smoothing times.

This alarm is automatically cleared when the average duration for calling the OBS metadata APIs is lower than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45175	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the average time for calling the OBS metadata APIs exceeds the threshold, the upper-layer big data computing services may be affected. To be more specific, the execution time of some computing tasks will exceed the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

Frame freezing occurs on the OBS server, or the network between the OBS client and the OBS server is unstable.

Handling Procedure

Check the heap memory usage.

Step 1 On the **FusionInsight Manager** homepage, choose **O&M** > **Alarm** > **Alarms** > **Average Time for Calling the OBS Metadata API Exceeds the Threshold**, view the role name in **Location**, and check the instance IP address.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Services** > **meta** > **Instance** > **meta** (IP address of the instance for which the alarm is generated). Click the drop-down list in the upper right corner of the chart area and choose **Customize**. In the dialog box that is displayed, select **Average time of OBS interface calls** from **OBS Meta data Operations**, and click **OK**. Check whether the average time of OBS metadata API calls exceeds the threshold.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **O&M** > **Alarm** > **Thresholds** > **meta** > **Average Time for Calling the OBS Metadata API**. Increase the threshold or smoothing times as required.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 In the **Services** area, select **NodeAgent**, **NodeMetricAgent**, **OmmServer**, and **OmmAgent** under OMS.

Step 7 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.306 ALM-45176 Success Rate of Calling OBS Metadata APIs Is Lower than the Threshold

Alarm Description

The system checks whether the success rate of calling OBS metadata APIs is lower than the threshold every 30 seconds. This alarm is generated when the success rate is lower than the threshold.

This alarm is automatically cleared when the success rate of calling APIs for writing OBS data is greater than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45176	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the success rate of calling the OBS metadata APIs is less than the threshold, the upper-layer big data computing services may be affected. To be more specific, some computing tasks may fail to be executed.

Possible Causes

An execution exception or severe timeout occurs on the OBS server.

Handling Procedure

Check the heap memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On the **FusionInsight Manager** homepage, choose **O&M** > **Alarm** > **Alarms** > **Success Rate for Calling the OBS Metadata API Is Lower Than the Threshold**, view the role name in **Location**, and check the instance IP address.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Services** > **meta** > **Instance** > **meta** (IP address of the instance for which the alarm is generated). Click the drop-down list in the upper right corner of the chart area and choose **Customize**. In the dialog box that is displayed, select **Success percent of OBS interface calls** from **OBS Meta data Operations**, and click **OK**. Check whether the average time of OBS metadata API calls exceeds the threshold.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **O&M** > **Alarm** > **Thresholds** > **meta** > **Success Rate for Calling the OBS Metadata API**. Increase the threshold or smoothing times as required.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 In the **Services** area, select **NodeAgent**, **NodeMetricAgent**, **OmmServer**, and **OmmAgent** under OMS.

Step 7 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.307 ALM-45177 Success Rate of Calling OBS Data Read APIs Is Lower than the Threshold

Alarm Description

The system checks whether the success rate of calling APIs for reading OBS data is lower than the threshold every 30 seconds. This alarm is generated when the success rate is lower than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is automatically cleared when the success rate of calling APIs for reading OBS data is greater than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45177	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the success rate of calling the OBS APIs for reading data is less than the threshold, the upper-layer big data computing services may be affected. To be more specific, some computing tasks may fail to be executed.

Possible Causes

An execution exception or severe timeout occurs on the OBS server.

Handling Procedure

Check the heap memory usage.

Step 1 On the **FusionInsight Manager** homepage, choose **O&M** > **Alarm** > **Alarms** > **Success Rate for Calling the OBS Data Read API Is Lower Than the Threshold**, view the role name in **Location**, and check the instance IP address.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Services** > **meta** > **Instance** > **meta** (IP address of the instance for which the alarm is generated). Click the drop-down list in the upper right corner of the chart area and choose **Customize**. In the dialog box that is displayed, select **Success percent of OBS data read operation interface calls** from **OBS data read operation**, and click **OK**. Check whether the average time of OBS metadata API calls exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **O&M** > **Alarm** > **Thresholds** > **meta** > **Success Rate for Calling the OBS Data Read API**. Increase the threshold or smoothing times as required.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 In the **Services** area, select **NodeAgent**, **NodeMetricAgent**, **OmmServer**, and **OmmAgent** under OMS.

Step 7 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.308 ALM-45178 Success Rate of Calling OBS Data Write APIs Is Lower Than the Threshold

Alarm Description

The system checks whether the success rate of calling APIs for writing OBS data is lower than the threshold every 30 seconds. This alarm is generated when the success rate is lower than the threshold.

This alarm is automatically cleared when the success rate of calling APIs for writing OBS data is greater than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45178	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the success rate of calling the OBS APIs for writing data is lower than the threshold, the upper-layer big data computing services may be affected. To be more specific, some computing tasks may fail to be executed.

Possible Causes

An execution exception or severe timeout occurs on the OBS server.

Handling Procedure

Check the heap memory usage.

Step 1 On the **FusionInsight Manager** homepage, choose **O&M** > **Alarm** > **Alarms** > **Success Rate for Calling the OBS Data Write API Is Lower Than the Threshold**, view the role name in **Location**, and check the instance IP address.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Services** > **meta** > **Instance** > **meta** (IP address of the instance for which the alarm is generated). Click the drop-down list in the upper right corner of the chart area and choose **Customize**. In the dialog box that is displayed, select **Success percent of OBS data write operation interface calls** from **OBS data write operation**, and click **OK**. Check whether the average time of OBS metadata API calls exceeds the threshold.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **O&M** > **Alarm** > **Thresholds** > **meta** > **Success Rate for Calling the OBS Data Write API**. Increase the threshold or smoothing times as required.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 In the **Services** area, select **NodeAgent**, **NodeMetricAgent**, **OmmServer**, and **OmmAgent** under OMS.

Step 7 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.309 ALM-45179 Number of Failed OBS readFully API Calls Exceeds the Threshold

Alarm Description

The system checks whether the number of failed OBS readFully API calls exceeds the threshold every 30 seconds. This alarm is generated when the number of failed API calls exceeds the threshold.

This alarm is automatically cleared when the number of failed OBS readFully API calls is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45179	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Certain upper-layer big data computing tasks will fail to execute.

Possible Causes

An execution exception or severe timeout occurs on the OBS server.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds**. On the **Thresholds** page, choose **meta** > **Number of failed calls to the OBS readFully interface**. In the right pane, set **Threshold** or **Trigger Count** to a larger value as required.

Step 2 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

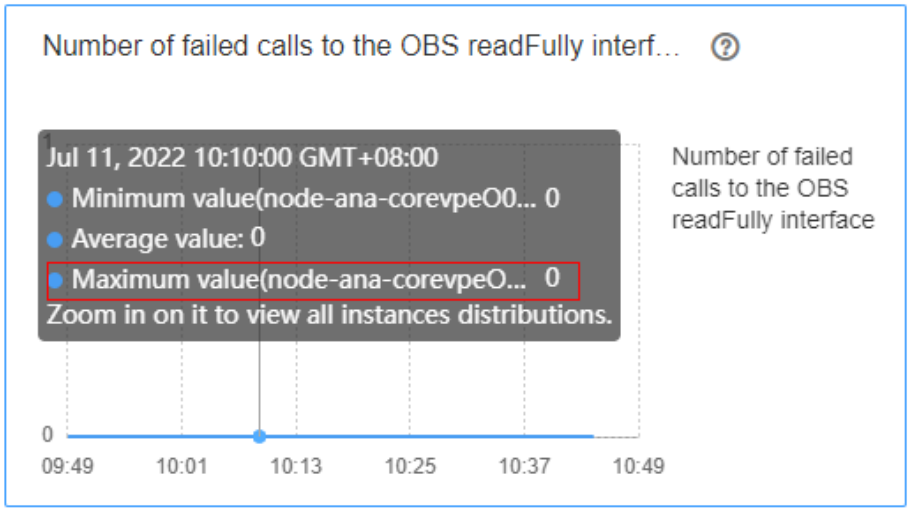
Step 3 Contact OBS O&M personnel to check whether the OBS service is normal.

- If yes, go to **Step 4**.
- If no, contact OBS O&M personnel to restore the OBS service.

Collect fault information.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **meta**. On the page that is displayed, click the **Chart** tab. On this tab page, select **OBS data read operation** in the **Chart Category** area. In the **Number of failed calls to the OBS readFully interface-All Instances** chart, view the host name of the instance that has the maximum number of failed OBS readFully API calls. For example, the host name is **node-ana-corevpeO003**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



Step 5 Choose **O&M** > **Log** > **Download** and select **meta** and **meta** under it for **Service**.

Step 6 Select the host obtained in **Step 4** for **Hosts**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.310 ALM-45180 Number of Failed OBS read API Calls Exceeds the Threshold

Alarm Description

The system checks whether the number of failed OBS read API calls exceeds the threshold every 30 seconds. This alarm is generated when the number of failed API calls exceeds the threshold.

This alarm is automatically cleared when the number of failed OBS read API calls is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45180	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Certain upper-layer big data computing tasks will fail to execute.

Possible Causes

An execution exception or severe timeout occurs on the OBS server.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds**. On the **Thresholds** page, choose **meta** > **Number of failed calls to the OBS read interface**. In the right pane, set **Threshold** or **Trigger Count** to a larger value as required.

Step 2 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

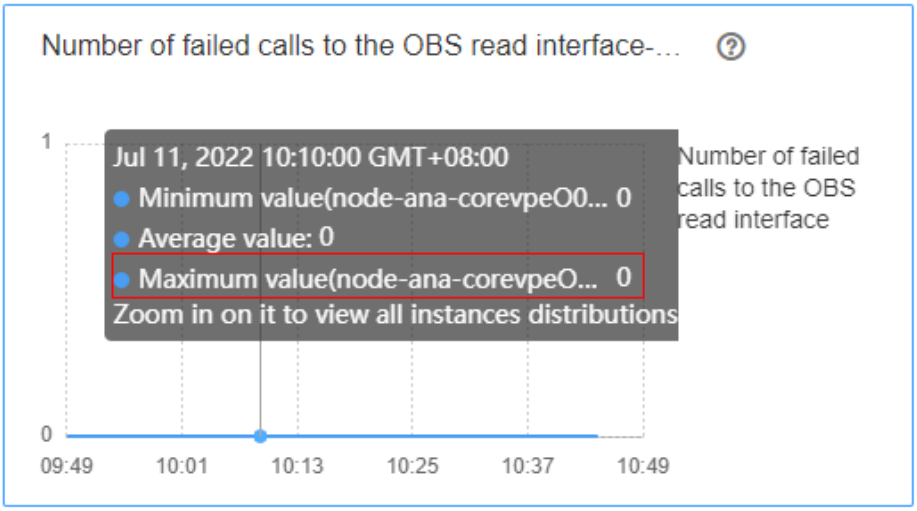
Step 3 Contact OBS O&M personnel to check whether the OBS service is normal.

- If yes, go to **Step 4**.
- If no, contact OBS O&M personnel to restore the OBS service.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **meta**. On the page that is displayed, click the **Chart** tab. On this tab page, select **OBS data read operation** in the **Chart Category** area. In the **Number of failed calls to the OBS read interface-All Instances** chart, view the host name of the instance that has the maximum number of failed OBS read API calls. For example, the host name is **node-ana-corevpeO003**.



Step 5 Choose **O&M** > **Log** > **Download** and select **meta** and **meta** under it for **Service**.

Step 6 Select the host obtained in **Step 4** for **Hosts**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.311 ALM-45181 Number of Failed OBS write API Calls Exceeds the Threshold

Alarm Description

The system checks whether the number of failed OBS write API calls exceeds the threshold every 30 seconds. This alarm is generated when the number of failed API calls exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is automatically cleared when the number of failed OBS write API calls is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45181	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Certain upper-layer big data computing tasks will fail to execute.

Possible Causes

An execution exception or severe timeout occurs on the OBS server.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds**. On the **Thresholds** page, choose **meta** > **Number of failed calls to the OBS write interface**. In the right pane, set **Threshold** or **Trigger Count** to a larger value as required.

Step 2 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

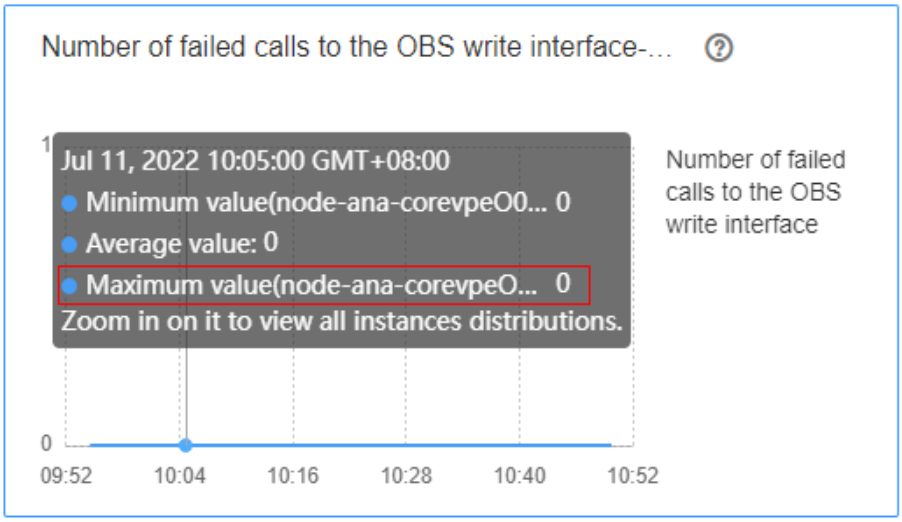
Step 3 Contact OBS O&M personnel to check whether the OBS service is normal.

- If yes, go to **Step 4**.
- If no, contact OBS O&M personnel to restore the OBS service.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Collect fault information.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **meta**. On the page that is displayed, click the **Chart** tab. On this tab page, select **OBS data write operation** in the **Chart Category** area. In the **Number of failed calls to the OBS write interface-All Instances** chart, view the host name of the instance that has the maximum number of failed OBS write API calls. For example, the host name is **node-ana-corevpeO003**.



Step 5 Choose **O&M** > **Log** > **Download** and select **meta** and **meta** under it for **Service**.

Step 6 Select the host obtained in **Step 4** for **Hosts**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.312 ALM-45182 Number of Throttled OBS Operations Exceeds the Threshold

Alarm Description

The system checks whether the number of throttled OBS operations exceeds the threshold every 30 seconds. This alarm is generated when the number of throttled OBS operations exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is automatically cleared when the number of throttled OBS operations is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45182	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Certain upper-layer big data computing tasks will fail to execute.

Possible Causes

The frequency of requesting OBS APIs is too high.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Thresholds**. On the **Thresholds** page, choose **meta** > **Number of Throttled OBS Operations**. In the right pane, set **Threshold** or **Trigger Count** to a larger value as required.

Step 2 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

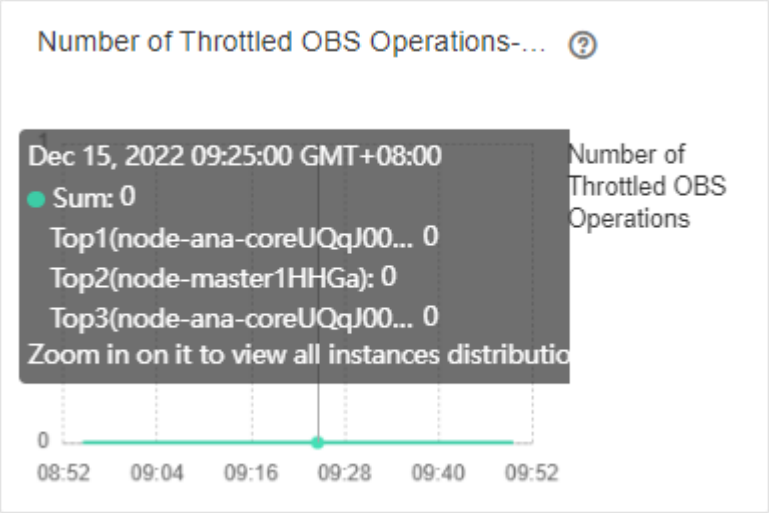
Step 3 Contact OBS O&M personnel to check whether the OBS service is normal.

- If yes, go to **Step 4**.
- If no, contact OBS O&M personnel to restore the OBS service.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **meta**. On the page that is displayed, click the **Chart** tab. On this tab page, select **OBS Throttled** in the **Chart Category** area. In the **Number of Throttled OBS Operations-All Instances** chart, view the host name of the instance that has the maximum number of throttled OBS API calls. For example, the host name is **node-ana-coreUQqJ0002**.



Step 5 Choose **O&M** > **Log** > **Download** and select **meta** and **meta** under it for **Service**.

Step 6 Select the host obtained in **Step 4** for **Hosts**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.313 ALM-45275 Ranger Service Unavailable

Alarm Description

The alarm module checks the Ranger service status every 180 seconds. This alarm is generated if the Ranger service is abnormal.

This alarm is cleared after the Ranger service recovers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45275	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

When the Ranger service is unavailable, Ranger cannot work properly and the native Ranger UI cannot be accessed.

Possible Causes

- The DBService service on which Ranger depends is abnormal.
- The RangerAdmin role instance is abnormal.

Handling Procedure

Check the DBService process status.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed page, check whether the ALM-27001 DBService Service Unavailable alarm is reported.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Rectify the DBService service fault by following the handling procedure of **ALM-27001 DBService Service Unavailable**. After the DBService alarm is cleared, check whether the "Ranger Service Unavailable" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check all RangerAdmin instances.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Log in to the node where the RangerAdmin instance is located as user **omm** and run the **ps -ef|grep "proc_rangeradmin"** command to check whether the RangerAdmin process exists on the current node.

- If yes, go to **Step 5**.
- If no, restart the faulty RangerAdmin instance or Ranger service and go to **Step 4**.

> **NOTICE**
>
> During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Step 4 In the alarm list, check whether the alarm "Ranger Service Unavailable" is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the fault that triggers the alarm is rectified, the alarm is automatically cleared.

Related Information

None

9.11.314 ALM-45276 Abnormal RangerAdmin Status

Alarm Description

The alarm module checks the RangerAdmin service status every 60 seconds. This alarm is generated if RangerAdmin is unavailable.

This alarm is automatically cleared after the RangerAdmin service recovers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45276	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Role for which the alarm is generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

If the status of a RangerAdmin is abnormal, access to the Ranger native UI is not affected. If there are two abnormal RangerAdmin instances, the Ranger native UI cannot be accessed and operations such as creating, modifying, and deleting policies are unavailable.

Possible Causes

The RangerAdmin port is not started.

Handling Procedure

Check the port process.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated.

Step 2 Log in to the node where the faulty RangerAdmin instance is located as user **root**.

Switch to user **omm**:

```
su - omm
```

Run the following command to obtain the PID of the RangerAdmin process:

```
ps -ef|grep "proc_rangeradmin" | grep -v grep | awk -F ' ' '{print $2}'
```

Run the following command to check whether the RangerAdmin process monitors the port (21401 for a cluster in security mode and 21400 for a cluster in normal mode):

```
netstat -anp | grep Process ID | grep LISTEN
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, restart the faulty RangerAdmin instance or Ranger service and go to **Step 3**.

> **NOTICE**
>
> During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Step 3 In the alarm list, check whether the "Abnormal RangerAdmin Status" alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect the fault information.

Step 4 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the fault that triggers the alarm is rectified, the alarm is automatically cleared.

Related Information

None

9.11.315 ALM-45277 RangerAdmin Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the RangerAdmin service every 60 seconds. This alarm is generated when the system detects that the heap memory usage of the RangerAdmin instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times. This alarm is cleared when the heap memory usage is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45277	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Heap memory overflow may cause service breakdown.

Possible Causes

The heap memory usage of the RangerAdmin instance is high or the heap memory is improperly allocated.

Handling Procedure

Check the heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45277 RangerAdmin Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **RangerAdmin Heap Memory Usage**. Click **OK**.

Step 3 Check whether the heap memory used by RangerAdmin reaches the threshold (95% of the maximum heap memory by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **RangerAdmin** > **Instance Configuration**. Click **All Configurations**, and choose **RangerAdmin** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for RangerAdmin cannot meet the heap memory required by the RangerAdmin process. You are advised to check the heap memory usage of RangerAdmin and change the value of **-Xmx** in **GC_OPTS** to the twice of the heap memory used by RangerAdmin. The value can be changed based on the actual service scenario. For details, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.316 ALM-45278 RangerAdmin Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the RangerAdmin service every 60 seconds. This alarm is generated when the direct memory usage of the RangerAdmin instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the direct memory usage of RangerAdmin is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45278	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Direct memory overflow may cause service breakdown.

Possible Causes

The direct memory of the RangerAdmin instance is overused or the direct memory is inappropriately allocated. As a result, the memory usage exceeds the threshold.

Handling Procedure

Check the direct memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45278 RangerAdmin Direct Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **RangerAdmin Direct Memory Usage**. Click **OK**.

Step 3 Check whether the direct memory used by RangerAdmin reaches the threshold (80% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **RangerAdmin** > **Instance Configuration**. Click **All Configurations**, and choose **RangerAdmin** > **System**. Increase the value of **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the direct memory configured for RangerAdmin cannot meet the direct memory required by the RangerAdmin process. You are advised to check the direct memory usage of RangerAdmin and change the value of **-XX:MaxDirectMemorySize** in **GC_OPTS** to the twice of the direct memory used by RangerAdmin. You can change the value based on the actual service scenario. For details, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.317 ALM-45279 RangerAdmin Non Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the non-heap memory usage of the RangerAdmin service every 60 seconds. This alarm is generated when the non-heap memory usage of the RangerAdmin instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the non-heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45279	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Non-heap memory overflow may cause service breakdown.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The non-heap memory usage of the RangerAdmin instance is high or the non-heap memory is improperly allocated.

Handling Procedure

Check non-heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45279 RangerAdmin Non Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **RangerAdmin Non Heap Memory Usage**. Click **OK**.

Step 3 Check whether the non-heap memory used by RangerAdmin reaches the threshold (80% of the maximum non-heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **RangerAdmin** > **Instance Configuration**. Click **All Configurations**, and choose **RangerAdmin** > **System**. Set **-XX: MaxPermSize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the non-heap memory size configured for the RangerAdmin instance cannot meet the non-heap memory required by the RangerAdmin process. You are advised to change the value of **-XX:MaxPermSize** in **GC_OPTS** to the twice of the current non-heap memory usage or change the value based on the site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.318 ALM-45280 RangerAdmin GC Duration Exceeds the Threshold

Alarm Description

The system checks the GC duration of the RangerAdmin process every 60 seconds. This alarm is generated when the GC duration of the RangerAdmin process exceeds the threshold (12 seconds by default) for five consecutive times. This alarm is cleared when the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45280	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The RangerAdmin responds slowly.

Possible Causes

The heap memory of the RangerAdmin instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Handling Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45280 RangerAdmin GC Duration Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated and click the drop-down list in the upper right corner of the chart area. Choose **Customize** > **GC** > **RangerAdmin GC Duration**. Click **OK**.

Step 3 Check whether the GC duration of the RangerAdmin process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **RangerAdmin** > **Instance Configuration**. Click **All Configurations**, and choose **RangerAdmin** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for RangerAdmin cannot meet the heap memory required by the RangerAdmin process. You are advised to check the heap memory usage of RangerAdmin and change the value of **-Xmx** in **GC_OPTS** to the twice of the heap memory used by RangerAdmin. The value can be changed based on the actual service scenario. For details, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.319 ALM-45281 UserSync Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the UserSync service every 60 seconds. This alarm is generated when the system detects that the heap memory usage of the UserSync instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times. This alarm is cleared when the heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45281	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Heap memory overflow may cause service breakdown.

Possible Causes

The heap memory usage of the UserSync instance is high or the heap memory is improperly allocated.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45281 UserSync Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **UserSync Heap Memory Usage**. Click **OK**.

Step 3 Check whether the heap memory used by UserSync reaches the threshold (95% of the maximum heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **UserSync** > **Instance Configuration**. Click **All Configurations**, and choose **UserSync** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for UserSync cannot meet the heap memory required by the UserSync process. You are advised to change the **-Xmx** value of **GC_OPTS** to twice that of the heap memory used by UserSync. You can change the value based on the actual service scenario. For details about how to check the UserSync heap memory usage, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.320 ALM-45282 UserSync Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the UserSync service every 60 seconds. This alarm is generated when the direct memory usage of the UserSync instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the UserSync direct memory usage is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45282	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Direct memory overflow may cause service breakdown.

Possible Causes

The direct memory of the UserSync instance is overused or the direct memory is inappropriately allocated. As a result, the memory usage exceeds the threshold.

Handling Procedure

Check the direct memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45282 UserSync Direct Memory Usage Exceeds the Threshold**. Check the location information of the alarm. Check the name of the instance host for which the alarm is generated.

Step 2 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Ranger** > **Instances**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the **Chart** area and choose **Customize** > **CPU and Memory** > **UserSync Direct Memory Usage**. Click **OK**.

Step 3 Check whether the direct memory used by the UserSync reaches the threshold (80% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On the FusionInsight Manager homepage, choose **Cluster** > **Services** > **Ranger** > **Instances** > **UserSync** > **Instance Configuration**. Click **All Configurations**, and choose **UserSync** > **System**. Set **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter to a larger value based on service requirements and save the configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If this alarm is generated, the direct memory configured for UserSync cannot meet the direct memory required by the UserSync process. Check the direct memory usage of UserSync and change the value of -XX:MaxDirectMemorySize in GC_OPTS to the twice of the direct memory used by UserSync. You can change the value based on service requirements. For details, see Step 2.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

> **NOTICE**
>
> During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.321 ALM-45283 UserSync Non Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the non-heap memory usage of the UserSync service every 60 seconds. This alarm is generated when the non-heap memory usage of the UserSync instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the non-heap memory usage is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45283	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Non-heap memory overflow may cause service breakdown.

Possible Causes

The non-heap memory of the UserSync process is overused or the non-heap memory is inappropriately allocated.

Handling Procedure

Check non-heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45283 UserSync Non Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **UserSync Non Heap Memory Usage**. Click **OK**.

Step 3 Check whether the non-heap memory used by UserSync reaches the threshold (80% of the maximum non-heap memory by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **UserSync** > **Instance Configuration**. Click **All Configurations**, and choose **UserSync** > **System**. Set **-XX: MaxPermSize** in the **GC_OPTS** parameter to a larger value based on site requirements and click **Save** to save the configuration.

NOTE

If this alarm is generated, the non-heap memory size configured for the UserSync instance cannot meet the non-heap memory required by the UserSync process. You are advised to change the **-XX:MaxPermSize** value of **GC_OPTS** to twice that of the current non-heap memory size or change the value based on the site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.322 ALM-45284 UserSync Garbage Collection (GC) Time Exceeds the Threshold

Alarm Description

The system checks the GC duration of the UserSync process every 60 seconds. This alarm is generated when the GC duration of the UserSync process exceeds the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

threshold (12 seconds by default) for five consecutive times. This alarm is cleared when the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45284	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Threshold for triggering the alarm.

Impact on the System

UserSync responds slowly.

Possible Causes

The heap memory of the UserSync instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Handling Procedure

Check the GC time.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45284 UserSync Garbage Collection (GC) Time Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated and click the drop-down list in the upper right corner of the chart area. Choose **Customize** > **GC** > **UserSync GC Duration**. Click **OK**.

Step 3 Check whether the GC duration of the UserSync process collected every minute exceeds the threshold (12 seconds by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **UserSync** > **Instance Configuration**. Click **All Configurations**, and choose **UserSync** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for UserSync cannot meet the heap memory required by the UserSync process. You are advised to change the value of **GC_OPTS** to the twice that of the heap memory used by UserSync. You can change the value based on the actual service scenario. For details about how to check the UserSync heap memory usage, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the fault that triggers the alarm is rectified, the alarm is automatically cleared.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.323 ALM-45285 TagSync Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the TagSync service every 60 seconds. This alarm is generated when the heap memory usage of the TagSync instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times. This alarm is cleared when the heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45285	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Heap memory overflow may cause service breakdown.

Possible Causes

The heap memory usage of the TagSync instance is high or the heap memory is improperly allocated.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45285 TagSync Heap Memory Usage Exceeds the Threshold**. Check the location

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **TagSync Heap Memory Usage**. Click **OK**.

Step 3 Check whether the heap memory used by TagSync reaches the threshold (95% of the maximum heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **TagSync** > **Instance Configuration**. Click **All Configurations** and choose **TagSync** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for TagSync cannot meet the heap memory required by the TagSync process. You are advised to change the **-Xmx** value of **GC_OPTS** to twice that of the heap memory used by TagSync. You can change the value based on the actual service scenario. For details about how to check the TagSync heap memory usage, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.324 ALM-45286 TagSync Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the TagSync service every 60 seconds. This alarm is generated when the direct memory usage of the TagSync instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the TagSync direct memory usage is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45286	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Direct memory overflow may cause service breakdown.

Possible Causes

The direct memory of the TagSync instance is overused or the direct memory is inappropriately allocated. As a result, the memory usage exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the direct memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45286 TagSync Direct Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **TagSync Direct Memory Usage**. Click **OK**.

Step 3 Check whether the direct memory used by the TagSync reaches the threshold (80% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **TagSync** > **Instance Configuration**. Click **All Configurations** and choose **TagSync** > **System**. Increase the value of **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the direct memory configured for TagSync cannot meet the direct memory required by the TagSync process. You are advised to check the direct memory usage of TagSync and change the value of **-XX:MaxDirectMemorySize** in **GC_OPTS** to the twice of the direct memory used by TagSync. You can change the value based on the actual service scenario. For details, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.325 ALM-45287 TagSync Non Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the non-heap memory usage of the TagSync service every 60 seconds. This alarm is generated when the non-heap memory usage of the TagSync instance exceeds the threshold (80% of the maximum memory) for five consecutive times. This alarm is cleared when the non-heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45287	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Non-heap memory overflow may cause service breakdown.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The non-heap memory of the TagSync process is overused or the non-heap memory is inappropriately allocated.

Handling Procedure

Check non-heap memory usage.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45287 TagSync Non Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **TagSync Non Heap Memory Usage**. Click **OK**.

Step 3 Check whether the non-heap memory used by TagSync reaches the threshold (80% of the maximum non-heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **TagSync** > **Instance Configuration**. Click **All Configurations** and choose **TagSync** > **System**. Set **-XX: MaxPermSize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the non-heap memory size configured for the TagSync instance cannot meet the non-heap memory required by the TagSync process. You are advised to change the **-XX:MaxPermSize** value of **GC_OPTS** to twice that of the current non-heap memory size or change the value based on the site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.326 ALM-45288 TagSync Garbage Collection (GC) Time Exceeds the Threshold

Alarm Description

The system checks the GC duration of the TagSync process every 60 seconds. This alarm is generated when the GC duration of the TagSync process exceeds the threshold (12 seconds by default) for five consecutive times. This alarm is cleared when the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45288	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

TagSync responds slowly.

Possible Causes

The heap memory of the TagSync instance is overused or the heap memory is inappropriately allocated. As a result, GCs occur frequently.

Handling Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45288 TagSync Garbage Collection (GC) Time Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated and click the drop-down list in the upper right corner of the chart area. Choose **Customize** > **GC** > **TagSync GC Duration**. Click **OK**.

Step 3 Check whether the GC duration of the TagSync process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **TagSync** > **Instance Configuration**. Click **All Configurations** and choose **TagSync** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter based on the site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for TagSync cannot meet the heap memory required by the TagSync process. You are advised to change the **-Xmx** value of **GC_OPTS** to twice that of the heap memory used by TagSync. You can change the value based on the actual service scenario. For details about how to check the TagSync heap memory usage, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect the fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.327 ALM-45289 PolicySync Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the PolicySync service every 60 seconds. This alarm is generated when the heap memory usage of the PolicySync instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times. This alarm is cleared when the heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45289	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Heap memory overflow may cause service breakdown.

Possible Causes

The heap memory of the PolicySync instance is overused or the heap memory is inappropriately allocated.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45289 PolicySync Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **PolicySync Heap Memory Usage**. Click **OK**.

Step 3 Check whether the heap memory used by PolicySync reaches the threshold (95% of the maximum heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **PolicySync**. Click **Instance Configuration** and then **All Configurations**, and choose **PolicySync** > **System**. Set **-Xmx** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for PolicySync cannot meet the heap memory required by the PolicySync process. You are advised to change the value of **-Xmx** in **GC_OPTS** to twice that of the heap memory used by PolicySync. You can change the value based on the actual service scenario. Refer to **Step 2** to view the PolicySync heap memory usage.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.328 ALM-45290 PolicySync Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the PolicySync service every 60 seconds. This alarm is generated when the direct memory usage of the PolicySync instance exceeds the threshold (90% of the maximum memory) for five consecutive times. This alarm is cleared when the PolicySync direct memory usage is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45290	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Direct memory overflow may cause service breakdown.

Possible Causes

The direct memory of the PolicySync process is overused or the direct memory is inappropriately allocated.

Handling Procedure

Check the direct memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45290 PolicySync Direct Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **PolicySync Direct Memory Usage**. Click **OK**.

Step 3 Check whether the direct memory used by the PolicySync reaches the threshold (90% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **PolicySync**. Click **Instance Configuration** and then **All Configurations**, and choose **PolicySync** > **System**. Set **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

If this alarm is generated, the direct memory configured for PolicySync cannot meet the requirements of the PolicySync process. You are advised to check the direct memory usage of PolicySync and change the value of **-XX:MaxDirectMemorySize** in **GC_OPTS** to twice (depending on the actual scenario) of the direct memory used by PolicySync. For how to view the PolicySync direct memory usage, see **Step 2**.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.329 ALM-45291 PolicySync Non-Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the non-heap memory usage of the PolicySync service every 60 seconds. This alarm is generated when the non-heap memory usage of the PolicySync instance exceeds the threshold (90% of the maximum memory) for five consecutive times. This alarm is cleared when the non-heap memory usage is less than the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45291	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Non-heap memory overflow may cause service breakdown.

Possible Causes

The non-heap memory of the PolicySync instance is overused or the non-heap memory is inappropriately allocated.

Handling Procedure

Check non-heap memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45291 PolicySync Non-Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **PolicySync Non-Heap Memory Usage**. Click **OK**.

Step 3 Check whether the non-heap memory used by PolicySync reaches the threshold (90% of the maximum heap memory by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **PolicySync**. Click **Instance Configuration** and then **All Configurations**, and choose **PolicySync** > **System**. Set **-XX: MaxPermSize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the non-heap memory size configured for the PolicySync instance cannot meet the non-heap memory required by the PolicySync process. You are advised to change the value of **-XX:MaxPermSize** in **GC_OPTS** to twice that of the current non-heap memory size or change the value based on site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.330 ALM-45292 PolicySync GC Duration Exceeds the Threshold

Alarm Description

The system checks the GC duration of the PolicySync process every 60 seconds. This alarm is generated when the GC duration of the PolicySync process exceeds

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the threshold for five consecutive times. This alarm is cleared when the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45292	Critical (default threshold: 20,000 ms) Major (default threshold: 12,000 ms)	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

PolicySync responds slowly.

Possible Causes

The heap memory of the PolicySync process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC duration.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45292 PolicySync GC Duration Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**. Select the role corresponding to the host name of the instance for which the alarm is

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

generated and click the drop-down list in the upper right corner of the chart area. Choose **Customize** > **GC** > **PolicySync GC Duration**. Click **OK**.

Step 3 Check whether the GC duration of the PolicySync process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance** > **PolicySync**. Click **Instance Configuration** and then **All Configurations**, and choose **PolicySync** > **System**. Set **-Xmx** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for PolicySync cannot meet the heap memory required by the PolicySync process. You are advised to change the value of **-Xmx** in **GC_OPTS** to twice that of the heap memory used by PolicySync. You can change the value based on the actual service scenario. Refer to **Step 2** to view the PolicySync heap memory usage.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the service restart, the service is interrupted. During the instance restart, the instance is unavailable, and the task on that instance fails to be executed.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.331 ALM-45293 Ranger User Synchronization Exception

Alarm Description

The system checks synchronization status of the UserSync process every 5 minutes. This alarm is generated when a synchronization exception occurs. This alarm is cleared when user synchronization becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45293	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the trigger condition, that is, an exception occurs during Ranger user synchronization.

Impact on the System

Unsynchronized users cannot access the native Ranger page and set permission policies for other users. As a result, some users may fail to access services that require Ranger permissions.

Possible Causes

The RangerAdmin instance is abnormal.

The UserSync instance is abnormal.

The LDAP service is abnormal.

Handling Procedure

Check whether the UserSync is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45293 Ranger User Synchronization Exception**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger** > **Instance**, select the UserSync instance on the host for which the alarm is generated, and check whether the instance status is abnormal.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 On FusionInsight Manager, choose **Cluster** > **Services** > **Ranger**, click **Instances**, and click **UserSync**. On the displayed page, click **More** > **Restart Instance**, or restart the Ranger service.

Step 4 Check whether the alarm is cleared in 5 to 10 minutes.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the RangerAdmin is abnormal.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** to check whether alarm "ALM-45275 Ranger Service Unavailable" is reported.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Rectify the fault by following the handling procedure of "ALM-45275 Ranger Service Unavailable".

Step 7 Check whether the alarm is cleared in 5 to 10 minutes.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the LDAP service is abnormal.

Step 8 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** to check whether alarm "ALM-25000 LdapServer Service Unavailable" is reported.

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Clear the alarm according to the handling procedure of "ALM-25000 LdapServer Service Unavailable".

Step 10 Check whether the alarm is cleared in 5 to 10 minutes.

- If yes, no further action is required.
- If no, go to **Step 11**.

Collect fault information.

Step 11 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 12 Expand the **Service** drop-down list, and select **Ranger** for the target cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.332 ALM-45294 RangerKMS Process Is Abnormal

Alarm Description

The RangerKMS process checks the process status every 20 seconds. This alarm is generated when the process status is abnormal and does not recover for a long time.

This alarm is cleared when the process status recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45294	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the alarm triggering condition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the process status is abnormal, the process cannot provide services properly. As a result, the entire service may become abnormal.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and some processes are in the D state and Z state. The process may also be suspended and enter the T state.

Handling Procedure

Check whether the process is in the D, Z, or T state.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. Wait for about 10 minutes and check whether the alarm is automatically cleared.

- If the alarm is not in the list, no further action is required.
- If the alarm is in the list, view the alarm details and record the IP address of the host where the alarm is generated. Run the command in **Step 2**.

Step 2 Log in to the host where the alarm is generated as the **root** user and run the **su - omm** command to switch to the **omm** user.

Step 3 Run the following command to check the process state:

ps ww -eo stat,cmd| grep -w Dproc_rangerkms | grep -v grep | awk '{print$1}'

Step 4 Check whether the command output contains any abnormal state (D, Z, or T).

- If the output contains any abnormal state, go to **Step 5**.
- If the output does not contain abnormal states, go to **Step 7**.

Step 5 Switch to user **root** and run the **reboot** command to restart the host for which the alarm is generated. (Restarting a host is risky. Ensure that the service process is normal after the restart.)

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If the alarm is cleared, no further action is required.
- If the alarm fails to be cleared, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **Ranger** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.333 ALM-45425 ClickHouse Service Unavailable

Alarm Description

The alarm module checks the ClickHouse instance status every 60 seconds. This alarm is generated when the alarm module detects that all ClickHouse instances are abnormal.

This alarm is cleared when the system detects that any ClickHouse instance is restored and the alarm is cleared.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45425	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The ClickHouse service is abnormal. You cannot use FusionInsight Manager to perform cluster operations on the ClickHouse service. The ClickHouse service function is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The configuration information in the **metrika.xml** file in the component configuration directory of the faulty ClickHouse instance node is inconsistent with that of the corresponding ClickHouse instance in the ZooKeeper.

Handling Procedure

Check whether the configuration in metrika.xml of the ClickHouse instance is correct.

Step 1 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Instance**, and locate the abnormal ClickHouse instance based on the alarm information.

- If yes, go to **Step 2**.
- If no, go to **Step 9**.

Step 2 Log in to the host where the ClickHouse service is abnormal and ping the IP address of another normal ClickHouse instance node to check whether the network connection is normal.

- If yes, go to **Step 3**.
- If no, contact the network administrator to repair the network.

Step 3 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port** 9000

Run the following command to query the value of **macros.id**:

select substitution from system.macros where macro='id';

Step 4 Log in to the host where the ZooKeeper client is located and log in to the ZooKeeper client.

Switch to the client installation directory.

Example: **cd /opt/client**

Run the following command to configure environment variables:

source bigdata_env

Run the following command to authenticate the user (skip this step in common mode):

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

kinit *Component service user*

Run the following command to log in to the client tool:

zkCli.sh -server *Service IP address of the node where the ZooKeeper role instance is located*:*client port*

Step 5 Run the following command to check whether the ClickHouse cluster topology information can be obtained.

get /clickhouse/config/*value of* ***macros.id*** *in* **Step 3**/**metrika.xml**

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 Log in to the host where the ClickHouse instance is abnormal and go to the configuration directory of the ClickHouse instance.

cd ${BIGDATA_HOME}**/FusionInsight_ClickHouse_***Version***/**x_x_**ClickHouseServer/etc**

cat metrika.xml

Step 7 Check whether the cluster topology information on ZooKeeper obtained in **Step 5** is the same as that in the **metrika.xml** file in the component configuration directory in **Step 6**.

- If yes, check whether the alarm is cleared. If the alarm persists, go to **Step 9**.
- If no, go to **Step 8**.

Step 8 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**, click **More**, and select **Synchronize Configuration**. Then, check whether the service status is normal and whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect the fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 11 Choose the corresponding host form the host list.

Step 12 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.334 ALM-45426 ClickHouse Service Quantity Quota Usage in ZooKeeper Exceeds the Threshold

Alarm Description

The alarm module checks the quota usage of the ClickHouse service in the ZooKeeper every 60 seconds. This alarm is generated when the alarm module detects that the usage exceeds the threshold (90%).

This alarm is cleared when the system detects that the usage is lower than the threshold and the alarm is cleared.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45426	Major (default)	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

After the ZooKeeper quantity quota of the ClickHouse service exceeds the threshold, you cannot perform cluster operations on the ClickHouse service on FusionInsight Manager. As a result, the ClickHouse service cannot be used.

Possible Causes

- When table data is created, inserted, or deleted, the ClickHouse creates znodes on ZooKeeper nodes. As the service volume increases, the number of znodes may exceed the configured threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- No quota limit is set for the metadata directory **/clickhouse** of ClickHouse in ZooKeeper.

Handling Procedure

Check the number of znodes created by ClickHouse on ZooKeeper.

Step 1 Log in to the host where the ZooKeeper client is located and log in to the ZooKeeper client.

Switch to the client installation directory.

Example: **cd /opt/client**

Run the following command to configure environment variables:

source bigdata_env

Run the following command to authenticate the user (skip this step in common mode):

kinit *Component service user*

Run the following command to log in to the client tool:

zkCli.sh -server *service IP address of the node where the ZooKeeper role instance locates*:*client port*

Step 2 Run the following command to check the quota used by the ClickHouse in the ZooKeeper and check whether the quota information is correctly set:

listquota /clickhouse

```
absolute path is /zookeeper/quota/clickhouse
Quota for path /clickhouse does not exist.
```

If the preceding information indicates that the quota configuration is incorrect, go to **Step 3**.

If no, go to **Step 5**.

Step 3 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **ZooKeeper**. On the displayed page, click **Configurations** and click **All Configurations**. On this sub-tab page, search for **quotas.auto.check.enable** to check whether its value is **true**.

If the value is not **true**, change the value to **true** and click **Save**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**, click **More**, and select **Synchronize Configuration**. After the synchronization is successful, go to **Step 1**.

Step 5 Run the following command and check whether the ratio of the **count** value of **Output stat** to the **count** value of **Output quota** in the command output is greater than **0.9**:

listquota /clickhouse

```
absolute path is /zookeeper/quota/clickhouse
Output quota for /clickhouse count=200000,bytes=1000000000
Output stat for /clickhouse count=2667,bytes=60063
```

In the preceding information, the **count** value of **Output stat** is **2667**, and the **count** value of **Output quota** is **200000**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 6**.
- If no, check whether the alarm is cleared 5 minutes later. If the alarm persists, go to **Step 8**.

Step 6 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**. Click **Configurations** and then **All Configurations**. Search for **clickhouse.zookeeper.quota.node.count**, set it to a value twice the **count** of **Output stat** in **Step 5**. **Do not use a value larger than 6000000. Otherwise, there will be high risks. Exercise caution when setting this parameter.**

Step 7 Restart the ClickHouse instance for which the alarm is generated, and check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, perform **Step 6** again, and check whether the alarm is cleared 5 minutes later. If the alarm persists, go to **Step 8**.

NOTICE

During the instance restart, the instance is unavailable, and the ClickHouse service on that instance fails to be executed.

Collect the fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 10 Choose the corresponding host form the host list.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.335 ALM-45427 ClickHouse Service Capacity Quota Usage in ZooKeeper Exceeds the Threshold

Alarm Description

The alarm module checks the quota usage of the ClickHouse service in the ZooKeeper every 60 seconds. This alarm is generated when the alarm module detects that the usage exceeds the threshold (90%).

This alarm is cleared when the system detects that the usage is lower than the threshold and the alarm is cleared.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45427	Major (default)	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

After the ZooKeeper quantity quota of the ClickHouse service exceeds the threshold, you cannot perform cluster operations on the ClickHouse service on FusionInsight Manager. As a result, the ClickHouse service cannot be used.

Possible Causes

- When table data is created, inserted, or deleted, the ClickHouse creates znodes on ZooKeeper nodes. As the service volume increases, the capacity of znodes may exceed the configured threshold.
- No quota limit is set for the metadata directory **/clickhouse** of ClickHouse in ZooKeeper.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the znode capacity of the ClickHouse in the ZooKeeper.

Step 1 Log in to the host where the ZooKeeper client is located and log in to the ZooKeeper client.

Switch to the client installation directory.

Example: **cd /opt/client**

Run the following command to configure environment variables:

source bigdata_env

Run the following command to authenticate the user (skip this step in common mode):

kinit *Component service user*

Run the following command to log in to the client tool:

zkCli.sh -server *Service IP address of the node where the ZooKeeper role instance is located*:*client port*

Step 2 Run the following command to check the quota used by the ClickHouse in the ZooKeeper and check whether the quota information is correctly set:

listquota /clickhouse

```
absolute path is /zookeeper/quota/clickhouse
Quota for path /clickhouse does not exist.
```

- If the preceding information indicates that the quota configuration is incorrect, go to **Step 3**.
- If not, go to **Step 5**.

Step 3 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **ZooKeeper**. On the displayed page, click **Configurations** and click **All Configurations**. On this sub-tab page, search for **quotas.auto.check.enable** to check whether its value is **true**.

If the value is not **true**, change the value to **true** and click **Save**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**, click **More**, and select **Synchronize Configuration**. After the synchronization is successful, go to **Step 1**.

Step 5 Run the following command and check whether the ratio of the **bytes** value of **Output stat** to the **bytes** value of **Output quota** in the command output is greater than **0.9**:

listquota /clickhouse

```
absolute path is /zookeeper/quota/clickhouse
Output quota for /clickhouse count=200000,bytes=1000000000
Output stat for /clickhouse count=2667,bytes=60063
```

In the preceding information, the **bytes** value of **Output stat** is **60063**, and the **bytes** value of **Output quota** is **1000000000**.

- If yes, go to **Step 6**.
- If no, check whether the alarm is cleared 5 minutes later. If the alarm persists, go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**. Click **Configurations** and then **All Configurations**. Search for **clickhouse.zookeeper.quota.size**, set it to a value twice the **bytes** of **Output stat** in **Step 5**. **Do not use a value larger than 6000000. Otherwise, there will be high risks. Exercise caution when setting this parameter.**

Step 7 Restart the ClickHouse instance for which the alarm is generated, and check whether the alarm is cleared 5 minutes later.

- If yes, no further action is required.
- If no, perform **Step 6** again, and check whether the alarm is cleared 5 minutes later. If the alarm persists, go to **Step 8**.

NOTICE

During the instance restart, the instance is unavailable, and the ClickHouse service on that instance fails to be executed.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 10 Choose the corresponding host form the host list.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.336 ALM-45428 ClickHouse Disk I/O Exception

Alarm Description

This alarm is generated when the alarm module detects EIO or EROFS errors during ClickHouse read and write every 60 seconds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45428	Major (default)	No

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

- ClickHouse fails to read and write data. The INSERT, SELECT, and CREATE operations on the local tables may be abnormal. Distributed tables are not affected.
- Services are affected, and I/Os fail.

Possible Causes

The disk is aged or has bad sectors.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45428 ClickHouse Disk I/O Exception**. Check the role name and the IP address of the host where the alarm is generated in **Location**.

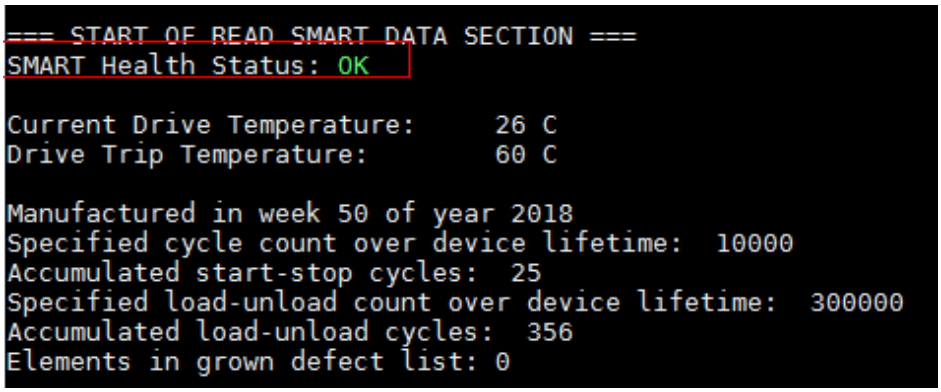
Step 2 Use PuTTY to log in to the node for which the fault is generated as user **root**.

Step 3 Run the **df -h** command to check the mount directory and find the disk mounted to the faulty directory.

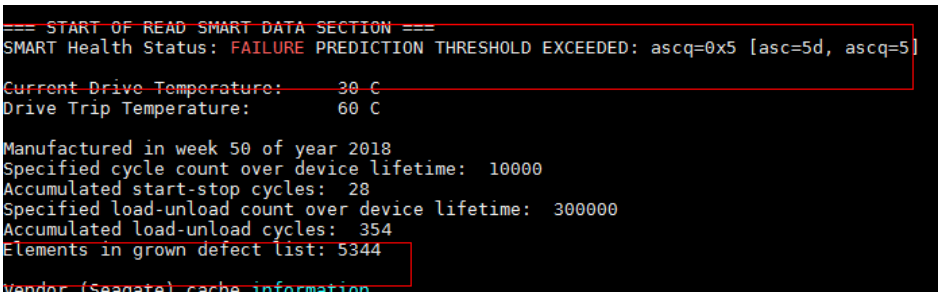
Step 4 Run the **smartctl -a /dev/sd***Faulty disk* command to check the disk. In the command, *Faulty disk* indicates the disk obtained in **Step 3**.

- If **SMART Health Status: OK** is displayed, as shown in the following figure, the disk is healthy. In this case, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



- If the number following **Elements in grown defect list** is not 0, as shown in the following figure, the disk may have bad sectors. If **SMART Health Status: FAILURE** is displayed, the disk is in the sub-health state. In this case, contact O&M personnel.



Step 5 After the fault is rectified, manually clear the alarm on FusionInsight Manager and check whether the alarm is generated again during the periodic check.

- If yes, go to **Step 6**.
- If no, no further action is required.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 8 Choose the corresponding host form the host list.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

If the alarm has no impact, manually clear the alarm.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.337 ALM-45429 Table Metadata Synchronization Failed on the Added ClickHouse Node

NOTE

This section applies only to MRS 3.1.2 or later.

Alarm Description

This alarm is generated when the local table corresponding to the distributed table fails to be created during ClickHouse capacity expansion.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45429	Major	No

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The distributed table fails to be queried.

Possible Causes

A node is stopped or faulty during capacity expansion.

Handling Procedure

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Instance**.

Step 2 Check whether an instance is stopped, decommissioned, or faulty.

- If yes, go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 4**.

Step 3 Start the instance or rectify the instance fault until all instances are running properly.

Step 4 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, locate this alarm and the faulty host based on the location information.

Step 5 Log in to the faulty host as user **omm**.

Step 6 Run the following commands to initialize environment variables:

source *Cluster installation directory*/**FusionInsight_ClickHouse_*/*_*_ClickHouseServer/etc/ENV_VARS**

source *Cluster installation directory*/**FusionInsight_ClickHouse_*/*_*_ClickHouseServer/etc/clickhouse-env.sh**

export CLICKHOUSE_CONF_DIR=${CLICKHOUSE_CONF_DIR}

Step 7 Run the following command to run the metadata synchronization tool to synchronize metadata from the existing node to the faulty node:

sh *Cluster installation directory*/**FusionInsight_ClickHouse_*/install/FusionInsight-ClickHouse-*/clickhouse/sbin/clickhouse-create-meta.sh true**

Step 8 Run the following command to view the log information and check whether the metadata has been synchronized:

vim /var/log/Bigdata/clickhouse/clickhouseServer/start.log

- If the synchronization is complete, go to **Step 9**.
- If the synchronization fails, go to **Step 10**.

Step 9 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Alarm ID** column, locate the corresponding alarm and click **Clear** in the **Operation** column. In the displayed dialog box, click **OK** to manually clear the alarm.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, select **ClickHouse** for the target cluster, and click **OK**.

Step 12 Choose the corresponding host form the host list.

Step 13 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 14 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm needs to be manually cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.338 ALM-45430 Permission Metadata Synchronization Failed on the Added ClickHouse Node

NOTE

This section applies only to MRS 3.1.2 or later.

Alarm Description

This alarm is generated when user and permission information fails to be synchronized during ClickHouse capacity expansion.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45430	Major	No

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The created user does not have operation permissions on the node.

Possible Causes

A node is stopped or faulty during capacity expansion.

Handling Procedure

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Instance**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Check whether an instance is stopped, decommissioned, or faulty.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Start the instance or rectify the instance fault until all instances are running properly.

Step 4 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, locate this alarm and the faulty host based on the location information.

Step 5 Log in to the faulty host as user **omm**.

Step 6 Run the following commands to initialize environment variables:

source ${BIGDATA_HOME}/FusionInsight_ClickHouse_*/*_*_ClickHouseServer/etc/ENV_VARS

source ${BIGDATA_HOME}/FusionInsight_ClickHouse_*/*_*_ClickHouseServer/etc/clickhouse-env.sh

export CLICKHOUSE_CONF_DIR=${CLICKHOUSE_CONF_DIR}

Step 7 Run the following command to run the metadata synchronization tool to synchronize metadata from the existing node to the faulty node:

sh ${BIGDATA_HOME}/FusionInsight_ClickHouse_*/install/FusionInsight-ClickHouse-*/clickhouse/sbin/clickhouse-create-meta.sh true

Step 8 Run the following command to view the log information and check whether the metadata has been synchronized:

vim /var/log/Bigdata/clickhouse/clickhouseServer/start.log

If the synchronization is complete, go to **Step 9**.

If the synchronization fails, go to **Step 10**.

Step 9 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the **Alarm ID** column, locate the corresponding alarm and click **Clear** in the **Operation** column. In the displayed dialog box, click **OK** to manually clear the alarm.

Collect the fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, select **ClickHouse** for the target cluster, and click **OK**.

Step 12 Choose the corresponding host form the host list.

Step 13 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 14 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm needs to be manually cleared after the fault is rectified.

Related Information

None

9.11.339 ALM-45431 Improper ClickHouse Instance Distribution for Topology Allocation

Alarm Description

The ClickHouseServer instance distribution does not meet the topology allocation requirements.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45431	Critical	No

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Some ClickHouseServer instances are unavailable.

Possible Causes

During installation or capacity expansion, the number of instances or allocation mode does not meet the topology requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**, locate the row that contains the alarm, and analyze the cause based on **Location** and **Additional Information**.

Step 2 Handle the alarm based on the additional alarm information and handling method in the following table.

Additional Information	Remarks	Handling Method
n ClickHouseServer instances should be added to other AZ.	This alarm is generated when a single cluster is deployed in cross-AZ DR mode. The ClickHouseServer instance deployment does not meet the cross-AZ DR topology allocation requirements. As a result, some instances cannot work properly.	1. On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**, click the **Instance** tab, locate the row that contains the alarm, view the host name in **Location**, and find the AZ for which the alarm is generated in the AZ column based on the host name. 2. On the **Instance** page, click **Add Instance** to add *n* ClickHouseServer instances to other AZs except the AZ where the alarm is generated.
n ClickHouseServer instances should be added.	This alarm is generated when a non-single-cluster is deployed in the default deployment mode of cross-AZ DR. The number of ClickHouseServer instances in the cluster is less than an even number. As a result, some instances cannot work properly.	1. Determine the number (*n*) of ClickHouseServer instances to be added based on the alarm information. 2. On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**, click the **Instance** tab, and add *n* ClickHouseServer instances to the cluster.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm needs to be manually cleared after the fault is rectified.

Related Information

None

9.11.340 ALM-45432 ClickHouse User Synchronization Process Fails

Alarm Description

The system checks the status of the ClickHouse user role synchronization process every 5 minutes. This alarm is generated when the system detects that the ClickHouse user role synchronization process is faulty or the user role synchronization fails.

This alarm is automatically cleared when the ClickHouse user role synchronization process or function becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45432	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Some ClickHouseServer instances are unavailable.

Possible Causes

- The ClickHouse user role synchronization process is not started properly or exits abnormally.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The ClickHouse user role synchronization process fails to synchronize user role information because the LdapServer service is faulty.

Handling Procedure

Check whether the ClickHouse user role synchronization process is normal.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, search for **ALM-45432 ClickHouse User Synchronization Process Fails**.

Step 2 Check the host name and additional information in the alarm details.

- If the additional information is "Process clickhouse-ugsync is not exit.", go to **Step 3**.
- If the additional information is "Process clickhouse-ugsync sync user failed.", go to **Step 6**.

Step 3 Log in to the faulty host as user **omm** and run the following command to check whether the ClickHouse user role synchronization process is normal:

ps -ef | grep 'clickhouse-ugsync'

Abnormal result of the synchronization process:

```
[omm@server-2110081635-0001 ~]$ ps -ef | grep 'clickhouse-ugsync'
omm      20104 13146  0 15:57 pts/7    00:00:00 grep --color=auto clickhouse-ugsync
```

- If yes, the alarm is automatically cleared. If the alarm is cleared, no further action is required. If the alarm persists, go to **Step 8**.
- If no, go to **Step 4**.

Step 4 Log in to the faulty host as user **omm** and run the following command to check whether the crontab daemon task is correctly configured:

crontab -l

Normal setting of the crontab daemon task:

```
*/5 * * * * bash /xxxxx/clickhouse_ugsync_check.sh >/dev/null 2>&1
```

- If yes, check whether the alarm is cleared 5 minutes later. If the alarm is cleared, no further action is required. If the alarm persists, go to **Step 8**.
- If no, go to **Step 5**.

Step 5 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**. On the page that is displayed, click the **Instance** tab. On this tab page, find the abnormal ClickHouseServer instance based on the fault information, and restart it. Wait for 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During the instance restart, the instance is unavailable, and the ClickHouse service on that instance fails to be executed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the LdapServer service is normal.

Step 6 Log in to FusionInsight Manager, choose **Cluster** > **Services**, and check whether **Running Status** of LdapServer is **Normal**.

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

Step 7 Handle the LdapServer service unavailable alarm according to ALM-25000 LdapServer Service Unavailable.

After **Running Status** of LdapServer becomes **Normal**, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 9 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **ClickHouse** for the target cluster.

Step 10 Expand the **Hosts** list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.341 ALM-45433 ClickHouse AZ Topology Exception

Alarm Description

If the cross-AZ HA function is enabled for a cluster where ClickHouse has been deployed, the ClickHouse topology remains unchanged. This alarm is generated when the cross-AZ HA does not take effect if backup nodes of the same shard are in the same AZ.

This alarm is automatically cleared when the system detects that all shards meet the cross-AZ HA deployment requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45433	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The current deployment of the ClickHouse service does not support cross-AZ HA.

Possible Causes

After cross-AZ HA is enabled, all backup nodes of a shard are in the same AZ.

Handling Procedure

Modify the AZ of backup nodes.

Step 1 Log in to the node where the client is installed as the client installation user. Run the following command to switch to the client installation directory:

cd *{Client installation path}*

Step 2 Run the following command to configure environment variables:

source bigdata_env

Step 3 Run the following command to authenticate the user (skip this step in normal mode):

kinit *Component service user*

Step 4 Run the following command to log in to the client tool:

zkCli.sh -server *Service IP address of the node where the ZooKeeper instance resides*:*Client port*

Step 5 Run the following command to view the current topology:

get /clickhouse/topo

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

> **NOTE**
>
> If the ClickHouse is installed with multiple services, run the **get /clickhouse{-n}/topo** command. For example, if the ClickHouse-1 is installed, run the **get /clickhouse-1/topo** command.

```
[zk: 192.168.20.36:24002(CONNECTED) 0] get /clickhouse/topo

<topo>
 <mcluster>
  <shard id="14" index="1">
   <server id="15">
    <replica>1</replica>
    <az>AZ1</az>
    <host>192-168-20-205</host>
    <port>21427</port>
   </server>
   <server id="16">
    <replica>2</replica>
    <az>AZ1</az>
    <host>192-168-20-2205</host>
    <port>21427</port>
   </server>
  </shard>
 </mcluster>
</topo>
```

Step 6 Select a host from the desired shard and deploy the host in another AZ.

Step 7 Log in to FusionInsight Manager, click **Host**, select the host you have deployed in **Step 6** and choose **More** > **Reinstall** to reinstall the host.

Step 8 Choose **Cluster** > **Cross-AZ HA**, click **Configure AZ and Policy** and change the AZ information of the reinstalled host to the AZ planned in the **Step 6**.

Step 9 Wait for five minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 11 Expand the drop-down list next to the **Service** field. In the **Services** dialog box that is displayed, select **ClickHouseServer** for the target cluster.

Step 12 Expand the **Hosts** list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 13 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.342 ALM-45434 A Single Replica Exists in the ClickHouse Data Table

Description

This alarm is generated when a single replica is detected in a custom logical cluster after the custom logical cluster is enabled for ClickHouse.

This alarm is automatically cleared when the system detects that the custom logical cluster uses multiple replicas.

Attribute

Alarm ID	Alarm Severity	Auto Clear
45434	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If a hardware fault occurs, data cannot be restored.

Possible Causes

The **metrika.xml** file in the ClickHouse configuration directory contains single-replica configuration.

Procedure

Check whether the configuration in metrika.xml of the ClickHouse instance is correct.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 In the alarm list on FusionInsight Manager, locate the row that contains the alarm, and click ⌄ to view the name of the host for which the alarm is generated. On the **Hosts** page, view the host IP address based on the host name.

Step 2 Log in to the host where the ClickHouse instance is abnormal, go to the configuration directory of the ClickHouse instance, and run the following commands:

cd ${BIGDATA_HOME}**/FusionInsight_ClickHouse_***Version***/**x_x_**ClickHouseServer/etc**

cat metrika.xml

Step 3 View the number of shards in each custom logical cluster and check that a single replica exists. Then, go to **Step 4**.

NOTE

If a shard contains only one node, a single replica exists in a logical cluster, as shown in the following:

```
<shard>
 <internal_replication>true</internal_replication>
 <replica>
  <host>host-name1</host>
  <port>port</port>
  <user>clickhouse</user>
  <password/>
 </replica>
</shard>
```

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 6 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 7 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.343 ALM-45435 Inconsistent Metadata of ClickHouse Tables

Alarm Description

This alarm is generated when the metadata in a distributed table or in the local table of the distributed table has been inconsistent for 180 min.

This alarm is automatically cleared when the metadata in the distributed table or in the local table of the distributed table becomes consistent.

Metadata consistency includes:

- Consistent quantity, name, sequence, and type of each column in the table
- Consistent partition keys
- Consistent sorting keys
- Consistent primary keys
- Consistent sampling keys

NOTE

If this alarm exists, table metadata is inconsistent in the ClickHouse cluster to which the current node belongs. The inconsistency may be caused by multiple reasons, not limited to those mentioned in additional information.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45435	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Table	Specifies the database name and table name for which the alarm is generated.

Impact on the System

Subsequent operations such as INSERT and ALTER on the table may fail.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

Table metadata modification fails or is not executed on one or more ClickHouseServer nodes.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port** 9000

Step 3 Check whether any task is being executed for the table to which the alarm is generated.

Run the following command to check whether any SQL task is being executed:

select * from system.processes where current_database='*Database name***' and query like '%***Table name***%'**

Run the following command to check whether a mutation task is being executed:

select * from system.mutations where database=*'Database name'* **and table=***'Table name'***;**

- If the query result is empty, go to **Step 4**.
- If the query result contains error information, rectify the fault accordingly. If the fault cannot be rectified based on the error information, go to **Step 6**.
- If the query result contains information about an on-going task with no error, the SQL/mutation task is being executed.

  Wait for 5 minutes. If the alarm is cleared, no further action is required. If the alarm persists, go to **Step 4**.

Step 4 Modify the table structure, delete a table, or add a table based on service requirements until the table metadata of all nodes in the cluster is consistent. After 5 minutes, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 If the table has been deleted, manually clear the alarm and check whether the alarm is reported again.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 6**.
- If no, no further action is required.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 8 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.344 ALM-45436 Skew ClickHouse Table Data

Alarm Description

This alarm is generated when data skew occurs in the local table of a distributed table between ClickHouse nodes. This alarm is automatically cleared when data becomes balanced.

Data skew check method:

- If **min_table_check_data_bytes** is set to **0**, data skew check is disabled.
- If **min_table_check_data_bytes** is greater than **0**, data skew check is enabled.

After data skew check is enabled, if the data volume in a table is less than the **min_table_check_data_bytes** value, no alarm will be reported due to data skew. When the data volume is greater than the **min_table_check_data_bytes** value and the data volume difference between the same table on different nodes is greater than the percentage specified in **min_table_data_varies_rate**, data skew occurs and this alarm is reported.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45436	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Table	Specifies the database name and table name for which the alarm is generated.

Impact on the System

SQL execution efficiency may be lowered.

Possible Causes

The data write policy is improper, causing unbalanced data among nodes.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- Security mode (with Kerberos enabled):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

- Normal mode (with Kerberos disabled):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port** 9000

Step 3 View data distribution.

select FQDN(), database, table, sum(data_compressed_bytes) from clusterAllReplicas(*Name of the logical cluster***, system.parts) where database='***Database name***' and table='***Table name***' and active=1 group by (FQDN(), database, table);**

Step 4 Balance data with a few clicks or migrate data based on service requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 8 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.345 ALM-45437 Excessive Parts in the ClickHouse Table

Alarm Description

This alarm is generated when the number of parts exceeds the threshold specified by **part_num_threshold**.

This alarm is automatically cleared when the number of parts is less than the **part_num_threshold** value.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45437	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Table	Specifies the database name and table name for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Service errors may occur.

Possible Causes

The data distribution in the ClickHouse table is improper, or the background merge task is executed slowly.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- Security mode (with Kerberos enabled):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

- Normal mode (with Kerberos disabled):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port** 9000

Step 3 Run the following command to manually merge parts:

optimize table *Database name*.*Table name* **final;**

Step 4 Check whether the number of parts has decreased.

select FQDN(), database, table, count(1) from clusterAllReplicas(default_cluster, system.parts) where database='*Database name***' and table='***Table name***' and active=1 group by (FQDN(), database, table);**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. If the number of parts is less than the threshold, wait for 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5**.
2. If the number of parts does not decrease, check whether the partition key of the table is set properly. If the number of partitions is too large, rectify the service logic.
3. If the command output is empty, the table does not exist. This alarm is a historical alarm and can be ignored. Manually clear it.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 7 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.346 ALM-45438 ClickHouse Disk Usage Exceeds 80%

Alarm Description

The system checks the disk capacity of the ClickHouseServer node every 1 minute. This alarm is generated when the usage of the disk where the ClickHouse data directory or metadata directory resides exceeds 80%.

This alarm is automatically cleared when the usage of the disk where the ClickHouse data directory or metadata directory is located is lower than 80%.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45438	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DiskPath	Specifies the path of the disk for which the alarm is generated.

Impact on the System

The ClickHouse write operation may fail.

Possible Causes

The disk capacity of the ClickHouseServer node is too small.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Expand the disk capacity of the node for which the alarm is generated.

Step 3 Go to **Step 4** if the expansion fails or the alarm persists after the expansion.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 6 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.347 ALM-45439 ClickHouse Node Enters the Read-Only Mode

Alarm Description

The system checks the disk capacity of the ClickHouseServer node every 1 minute. This alarm is generated when the system detects that the disk capacity exceeds 90% and the ClickHouseServer node enters the read-only mode.

This alarm is automatically cleared when the system detects that the disk capacity is lower than 90% and the ClickHouseServer node exits the read-only mode.

NOTE

If the ClickHouseServer node is in read-only mode and you need to log in to the client to clear data, you can manually exit the read-only mode using the following method:

Log in to FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Configurations** > **All Configurations**, search for **profiles.default.readonly**, and change its value to **0**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45439	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DiskPath	Specifies the path of the disk for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

After the ClickHouseServer node enters the read-only mode, all write, modification, and deletion operations fail.

Possible Causes

The disk usage of the ClickHouse node exceeds 90%.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Expand the disk capacity of the node for which the alarm is generated.

Step 3 Go to **Step 4** if the expansion fails or the alarm persists after the expansion.

NOTE

After the capacity expansion, this alarm can be automatically cleared only when **profiles.default.readonly** is **auto**. If its value has been manually changed, change it back to **auto**. If **profiles.default.readonly** needs to be set to **0** or **1** based on service requirements, manually clear this alarm.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 6 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.348 ALM-45440 Inconsistency Between ClickHouse Replicas

Alarm Description

When the number of ClickHouse replicas is greater than 1, the system periodically checks the replicated table. This alarm is generated if replicated table data is not

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

synchronized. This alarm is cleared when data in all replicated tables between replicas becomes synchronized.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45440	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Table	Specifies the table name for which the alarm is generated.

Impact on the System

The data reliability of the ClickHouse replicated table is affected, causing data differences and affecting the query result of the distributed table.

Possible Causes

- The ClickHouse service is overloaded.
- The connection between the ClickHouse and ZooKeeper is abnormal.

Handling Procedure

Check whether the ClickHouse service load is heavy.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the database name, table name, role name and IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port** 9000

Step 3 Run the following statement to check whether data is frequently written to the system table. If yes, wait until the service execution is complete and check whether the alarm is cleared.

SELECT query_id, user, FQDN(), elapsed, query FROM system.processes ORDER BY query_id;

- If yes, no further action is required.
- If no, go to **Step 4**.

Step 4 Check whether a large amount of data is written. If yes, wait until the task is complete and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Run the following statement to check whether replicas are synchronized:

select table,absolute_delay, queue_size, inserts_in_queue, merges_in_queue from system.replicas where absolute_delay > 0 order by absolute_delay desc limit 10;

- If yes, go to **Step 6**.
- If no, go to **Step 9**.

Step 6 If **inserts_in_queue** contains a large amount of content to be inserted, run the following SQL statement to query the replica synchronization queue and locate the error cause:

SELECT database,table,type,any(last_exception),any(postpone_reason),min(create_time),max(last_attempt_time),max(last_postpone_time),max(num_postponed) AS max_postponed,max(num_tries) AS max_tries,min(num_tries) AS min_tries,countIf(last_exception != '') AS count_err,countIf(num_postponed > 0) AS count_postponed,countIf(is_currently_executing) AS count_executing,count() AS count_all FROM system.replication_queue GROUP BY database,table,type ORDER BY count_all DESC

Check whether an error message similar to the following is displayed:

```
Not executing fetch of part xxx because n fetches already executing, max n
```

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Configurations** > **All Configurations**, and check whether the value of **background_pool_size** is twice the number of cores on the node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 8 Set this parameter to twice the number of cores on the node and synchronize the configuration. Wait for a while and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check the connectivity between ClickHouse and ZooKeeper.

Step 9 Log in to the node where the ClickHouseServer instance is located, go to **${BIGDATA_HOME}/FusionInsight_ClickHouse_*/*_*_ClickHouseServer/etc**, and check whether the port numbers of the ClickHouseServer and ZooKeeper in the **config.xml** file are the same, as shown in the following information in bold:

NOTE

To view the ZooKeeper port number, choose **Cluster** > **Services** > **ZooKeeper** > **Configurations** > **All Configurations** on FusionInsight Manager, and check the value of **clientPort**.

```
<zookeeper>
  <session_timeout_ms>10000</session_timeout_ms>
  <node index="1">
    <host>server-2110082001-0019</host>
    <port>24002</port>
  </node>
  <node index="2">
    <host>server-2110082001-0018</host>
    <port>24002</port>
  </node>
  <node index="3">
    <host>server-2110082001-0017</host>
    <port>24002</port>
  </node>
</zookeeper>
```

- If yes, go to **Step 11**.
- If no, go to **Step 10**.

Step 10 Change the port number to the ZooKeeper port number, restart the ClickHouseServer instance, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

NOTICE

During the instance restart, the instance is unavailable, and the ClickHouse service on that instance fails to be executed.

Collect fault information.

Step 11 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 12 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 13 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 14 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 15 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.349 ALM-45441 Zookeeper Disconnected

Alarm Description

The system checks the connection between ClickHouse and ZooKeeper every minute. This alarm is generated when the connection fails. The alarm is reported because the ZooKeeper connection is abnormal. If the connection fails for three consecutive times, the system generates an alarm.

This alarm is automatically cleared when the system detects that the connection is normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45441	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If ClickHouse is disconnected from ZooKeeper, the ClickHouse service cannot be used.

Possible Causes

- The ZooKeeper service is abnormal.
- The ClickHouse service is overloaded.

Handling Procedure

Check whether ZooKeeper is normal.

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **ZooKeeper** > **quorumpeer**.

Step 2 Check whether ZooKeeper instances are normal.

- If yes, go to **Step 6**.
- If no, go to **Step 3**.

Step 3 Select instances whose status is not good and choose **More** > **Restart Instance**.

NOTICE

During the instance restart, the instance is unavailable, and the ZooKeeper service on the current instance node fails to be executed.

Step 4 Check whether the instance status is good after restart.

- If yes, go to **Step 5**.
- If no, go to **Step 10**.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the ClickHouse service load is heavy.

Step 6 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 7 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host IP address of the ClickHouseServer instance that reports the alarm **--user** *User name* **--password --port** 9440

Step 8 Run the following statement to check whether data is frequently written to the system table. If yes, wait until the service execution is complete and check whether the alarm is cleared.

SELECT query_id, user, FQDN(), elapsed, query FROM system.processes ORDER BY query_id;

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Check whether a large amount of data is written. If yes, wait until the task is complete and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 12 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 13 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.350 ALM-45442 Too Many Concurrent SQL Statements

Alarm Description

The alarm module checks the number of concurrent ClickHouse requests every 30 seconds. This alarm is generated when the number of concurrent ClickHouse requests exceeds the concurrency threshold configured on the UI.

This alarm is cleared when the system detects that the actual number of concurrent requests is less than concurrency threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45442	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If there are too many concurrent SQL statements, a large number of system resources are consumed. As a result, system response becomes slow.

Possible Causes

The ClickHouse service is overloaded.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Choose **Cluster** > **ClickHouse** > **Instance**, select an instance based on the alarm information. Choose **Chart** > **Concurrency** to check whether the actual number of concurrent SQL statements is greater than SQL concurrency threshold.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Confirm with the user whether a large number of tasks were being executed during the alarming period.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 On FusionInsight Manager, choose **O&M** and click **Alarm** > **Thresholds** in the navigation pane on the left. On the displayed page, click **ClickHouse** > **Concurrency** and adjust the threshold, or wait until the task is complete. Check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 7 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.351 ALM-45443 Slow SQL Queries in the Cluster

Alarm Description

The system checks slow SQL queries for ClickHouse every 1 minute. This alarm is generated when the execution time of a SQL statement is longer than or equal to the slow SQL threshold.

This alarm is automatically cleared when the system detects that the execution time of the SQL statement is shorter than the slow SQL threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45443	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The performance of the ClickHouse service deteriorates, which slows the response of other services. If there are too many slow SQL statements, the service may be unavailable.

Possible Causes

- The ClickHouse service is overloaded.
- The execution of SQL statements takes a long time.

Handling Procedure

Check whether the ClickHouse service load is heavy.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port --secure**

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port**

Step 3 Run the following statement to check whether data is frequently written to the system table. If yes, wait until the service execution is complete and check whether the alarm is cleared.

SELECT query_id, user, FQDN(), elapsed, query FROM system.processes ORDER BY query_id;

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 4**.

Checking whether the SQL statements take a long time.

Step 4 Check the logical cluster to which the alarm object belongs. Log in to FusionInsight Manager, click **Cluster**, choose **Services** > **ClickHouse**, and click **Logic Cluster**. On the displayed page, choose **Query Management** > **Ongoing Slow Queries**. Check which SQL statements take a long time on the displayed page, confirm with the user to adjust services, optimize slow SQL statements, and check whether the optimization is successful.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 After the SQL statements are complete, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 8 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.352 ALM-45444 Abnormal ClickHouse Process

Alarm Description

The health check module checks ClickHouse instances every 30 seconds. If the number of consecutive failures exceeds the threshold, an alarm is reported. In this case, the ClickHouse process may stop responding and services cannot be properly executed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45444	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the ClickHouse process is abnormal, services cannot run properly.

Possible Causes

The ClickHouse process runs improperly.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 9440 **--secure**

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--user** *Username* **--password --port** 9000

Step 3 Run the following statement to check whether the result can be properly returned:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

SELECT 1;

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Wait for several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 7 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.353 ALM-45445 Failed to Send Data Files to Remote Shards When ClickHouse Writes Data to a Distributed Table

NOTE

This section is available for MRS 3.3.1 or later version only.

Alarm Description

The ClickHouse instance checks the distributed table every 300 seconds. If the number of consecutive failures exceeds the threshold, an alarm is generated. In this case, the node where the ClickHouse instance writes data to the distributed table cannot send data files to remote shard nodes.

This alarm is cleared when the number of consecutive failures falls below the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45445	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

The results of operations such as distributed table queries are abnormal.

Possible Causes

The status of some ClickHouse shard nodes is abnormal.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address of the hostname in **Location**.

Step 2 Log in to the node where the client is installed as the client installation user and run the following commands:

cd *{Client installation path}*

source bigdata_env

- For a cluster with Kerberos authentication enabled (security mode):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance for which the alarm is reported* **--port** 9440 **--secure**

- For a cluster with Kerberos authentication disabled (normal mode):

  clickhouse client --host *IP address of the ClickHouseServer instance for which the alarm is reported* **--user** *Username* **--password --port** 9000

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Run the following SQL statement to obtain the shard based on the value of **data_path**. For example, if the value of **data_path** is **/srv/Bigdata/clickhouse/data1.../shard2_all_replicas**, the desired shard is **shard2**.

select database, table, data_path, data_files, error_count from system.distribution_queue where data_files != 0 and error_count != 0;

Step 4 Run the following SQL statement to obtain the node IP address (value of the **host** field in the system table **system.clusters**) of the shard where data fails to be sent to (**shard_num** obtained in **Step 3**):

select * from system.clusters;

Step 5 Log in to the ClickHouse node (IP address obtained in **Step 2**) by referring to **Step 4**, run the following statement, and check whether the result can be returned:

SELECT 1;

If yes, go to **Step 6**.

If no, go to **Step 10**.

Step 6 Log in to the ClickHouse node (IP address obtained in **Step 2**) by referring to **Step 4**, run the following statement (**database_name** and **table_name** indicate the database name and table name of the local table corresponding to the distributed table):

select name,type from system.columns where database='*database_name***' and table='***table_name***'**

Step 7 Log in to the ClickHouse node for which the alarm is generated by referring to **Step 2** and run the following statement (**database_name** and **table_name** indicate the database name and table name of the distributed table to which data is written):

select name,type from system.columns where database='*database_name***' and table='***table_name***'**

Step 8 Check whether the results obtained in **Step 6** and **Step 7** are the same.

If yes, go to **Step 10**.

If no, ensure that the column information of the distributed table is consistent with that of the corresponding local table, and then try to write data to the distributed table.

Step 9 Wait for several minutes and check whether the alarm is cleared.

If yes, no further action is required.

If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 12 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 14 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.354 ALM-45446 Mutation Task of ClickHouse Is Not Complete for a Long Time

NOTE

This section is available for MRS 3.3.1 or later version only.

Alarm Description

The system checks mutation tasks every 5 minutes. This alarm is generated when the system detects that a mutation task has been running for at least **slow_mutation_cost_time** minutes. This alarm is automatically cleared when the system does not detect any running mutation task or the running time of a mutation task is less than **slow_mutation_cost_time** minutes.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45446	Minor	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

- Server resources are occupied, and the performance of the ClickHouse service deteriorates.
- Data is inconsistent.

Possible Causes

The data volume is too large. As a result, the mutation task runs slowly or is suspended.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Log in to the node where the client is installed and run the following commands:

cd *{Client installation path}*

source bigdata_env

- Security mode (with Kerberos enabled):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance for which the alarm is reported* **--port** 21427 **--secure**

- Normal mode (with Kerberos disabled):

  clickhouse client --host *IP address of the ClickHouseServer instance for which the alarm is reported* **--user** *Username* **--password --port** 21423

Step 3 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse**, click **Configurations** and then **All Configurations**. Search for the value of the **slow_mutation_cost_time** parameter, enter the parameter value in the following SQL statement, and run the following statement to check whether any result is returned:

SELECT * FROM system.mutations WHERE is_done = 0 AND create_time < now() - INTERVAL *The value* **SECOND**

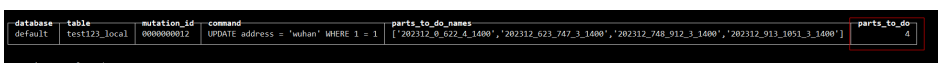
NOTE

Add the actual value of **slow_mutation_cost_time** to the preceding statement.

- If yes, go to **Step 4**.
- If no, go to **Step 7**.

Step 4 Wait for a while and run the statement in **Step 3** again. Check whether the value of **parts_to_do** in the returned result decreases.

- If yes, wait until the mutation task is complete.
- If no, go to **Step 5**.

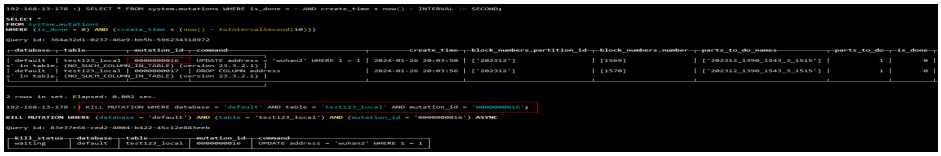


 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 If the value of **parts_to_do** remains unchanged, stop the mutation task. Run the following statement and run the statement in **Step 3** again to check whether the current mutation task is in the returned result list:

KILL MUTATION WHERE database = '*Database name***' AND table = '***Table name***' AND mutation_id ='mutation ID'**

- If yes, go to **Step 7**.
- If no, go to **Step 6**.



Step 6 Wait for several minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 9 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.355 ALM-45447 ClickHouse Table Read-Only

NOTE

This section applies only to MRS 3.3.1 or later.

Alarm Description

The system checks the table status every minute. This alarm is generated when the system detects that a table is read-only. This alarm is automatically cleared when no table is read-only.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45447	MRS 3.3.1-LTS and MRS 3.5.0-LTS: Major MRS 3.5.0-LTS.x and MRS 3.5.0-LTS or later: Minor	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

- Data cannot be written to or modified.
- Data synchronization in the replication table is interrupted, causing data inconsistency.

Possible Causes

The ZooKeeper is overloaded and metadata is lost.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address of the hostname in **Location**.

Step 2 Log in to the node where the client is installed and run the following commands:

cd *{Client installation path}*

source bigdata_env

- Security mode (with Kerberos enabled):

  kinit *Component service user*

  clickhouse client --host *IP address of the ClickHouseServer instance that reports the alarm* **--port** 21427 **--secure**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Normal mode (with Kerberos disabled):

  clickhouse client --host *IP address of the ClickHouseServer instance for which the alarm is reported* **--user** *Username* **--password --port** 21423

Step 3 Run the following SQL statement to check whether any table is in the read-only state:

select database,table from system.replicas where is_readonly = 1

- If yes, go to **Step 4**.
- If no, go to **Step 8**.

Step 4 Specify the database and table queried in **Step 3** in the following SQL statements and run them in sequence. Then, run the SQL statement in **Step 3** and check whether the result contains the read-only table.

system restore replica *database.table*;

system restart replica *database.table*;

- If yes, go to **Step 5**.
- If no, go to **Step 8**.

Step 5 Specify the database and table queried in **Step 3** in the following SQL statements and run them in sequence. Then, run the SQL statement in **Step 3** and check whether the result contains the read-only table.

detach table *database.table*;

attach table *database.table*;

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

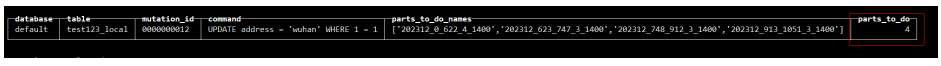
Step 6 Run the following SQL statement to view the structure information of the read-only table. In the statement, database and table are those queried in **Step 3**.

show create table *database.table*;

Step 7 Run the following SQL statement to delete the read-only table and create a read-only table based on the table structure information in **Step 6**. Wait for several minutes, run the SQL statement in **Step 3**, and check whether the result contains the read-only table.

drop table database.table no delay;

- If yes, go to **Step 9**.
- If no, go to **Step 8**.



Step 8 Wait several minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 11 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 12 Click the edit icon in the upper right corner and select a time span starting 1 hour before and ending 1 hour after when the alarm was generated. Then, click **Download** to collect the logs.

Step 13 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.356 ALM-45448 Rapid Increase of Znodes Used by ClickHouse

NOTE

This section is available for MRS 3.3.1 or later version only.

Alarm Description

Metadata in ClickHouse is stored on ZooKeeper Znodes. The number of occupied Znodes may increase sharply even if service requirements do not change greatly. This alarm is generated when the number of occupied Znodes in two hours exceeds the threshold. If a large amount of data is imported or services are migrated, ignore this alarm.

This alarm is cleared when the system detects that the increase in two hours is lower than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45448	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

ZooKeeper server resources are occupied. The number of Znodes will reach the upper limit in a short period of time, affecting the ClickHouse service.

Possible Causes

- If the ClickHouse service volume retains at large scale, increase the value of **ZNODE_GROWTH_LIMIT** on the ClickHouse configuration page.
- Major service changes, migration, and data import have been performed in ClickHouse.

Handling Procedure

Step 1 Choose **Cluster** > **ClickHouse** > **Instances** > **All Configurations**, and search for **ZNODE_GROWTH_LIMIT**, increase the value of **ZNODE_GROWTH_LIMIT**. The default value is **100000**, the maximum value is **200000**, and the value is configured based a step of 50000. Wait for 2 hours and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 2**.

Step 2 Confirm with the service party whether there are new service requests or a large amount of data is imported or migrated during the period when the alarm is reported.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Wait for 2 hours and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 6 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.357 ALM-45449 The Counter Number of zxid Used by ClickHouse Exceeds the Threshold

NOTE

This section is available for MRS 3.3.1 or later version only.

Alarm Description

ClickHouse uses ZooKeeper Transaction ID (**zxid**) to manage transactions. The **zxid** is a 64-bit number used to keep consistency of the distributed system. The **zxid** has two parts: the high order 32-bits for the epoch and the low order 32-bits for the counter. The epoch number represents the lifecycle of the leader, and the counter number indicates the location of a transaction in the request. Each time a new transaction is generated, the counter number is automatically increased by 1. When the value of **zxid** reaches the maximum value, that is, the counter number reaches **0xffffffff**, the cluster forcibly elects the leader and ZooKeeper is unavailable in a short period. The system checks the counter number every two hours. This alarm is generated when the counter number of **zxid** exceeds the threshold.

This alarm is cleared when the counter number is smaller than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45449	Warning	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

ZooKeeper leader election is forcibly triggered at an uncertain time, which may interrupt the ClickHouse service.

Possible Causes

The counter number of ZooKeeper zxid exceeds the threshold.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **ZooKeeper**. On the **Dashboard** page of the ZooKeeper service, click **More** and select **Service Rolling Restart** in the upper right corner. In the displayed dialog box, enter the password and click **OK**. On the **Service Rolling Restart** page, click **OK** and wait until the rolling restart of the ZooKeeper service is complete.

NOTE

Restart the ZooKeeper service during off-peak hours of ClickHouse.

Step 2 Wait for 2 hours and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Collect fault information.

Step 3 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 4 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 5 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.358 ALM-45450 ClickHouse Failed to Obtain a Temporary Agency Credential

NOTE

This section is available for MRS 3.3.1 or later version only.

Alarm Description

After the cold-hot separation function and an agency are configured, the system checks the status of the temporary agency credential every minute. This alarm is generated when the temporary agency credential fails to be obtained for three consecutive times.

This alarm is automatically cleared when the system successfully obtains the temporary agency credential.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45450	Critical	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The system cannot access OBS after the temporary agency credential expires. Operations such as reading and writing cold data in OBS cannot be performed on tables configured with cold and hot separation policies.

Possible Causes

- The OBS parameters configured for ClickHouse are incorrect.
- The IAM service is abnormal.

Handling Procedure

Step 1 Check whether the cold and hot separation configuration is correct. If the configuration is incorrect, modify the configuration and restart the ClickHouse instance. Wait for 3 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 2**.

Collect fault information.

Step 2 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 3 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 4 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 5 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 6 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.359 ALM-45451 ClickHouse Failed to Access OBS

NOTE

This section is available for MRS 3.3.1 or later version only.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Description

When cold-hot separation is enabled, the system checks the OBS access every minute. This alarm is generated when the system detects that OBS cannot be accessed for three consecutive times.

This alarm is automatically cleared when the system successfully accesses OBS.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45451	Critical	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

For tables configured with the cold-hot separation policy, cold data on OBS cannot be read or written. Hot data on local disks cannot be moved to OBS.

Possible Causes

- Parameters such as the endpoint used by ClickHouse to access OBS are incorrect.
- OBS is abnormal.

Handling Procedure

Step 1 Check whether the cold and hot separation configuration is correct. If the configuration is incorrect, modify the configuration and restart the ClickHouse instance. Wait for 3 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 2**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 3 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 4 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 5 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 6 Contact O&M personnel and provide the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.360 ALM-45452 ClickHouse's Local Disk Space Is Below the Cold-Hot Separation Threshold

NOTE

This section is available for MRS 3.3.1 or later version only.

Alarm Description

When cold-hot separation is enabled, the system checks the remaining space of the local disk specified in the cold-hot separation policy every 5 minutes. This alarm is generated if the remaining space is lower than the **move_factor** threshold.

This alarm is automatically cleared when the remaining space of the local disk is greater than the **move_factor** threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45452	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

Some hot data on the local disk is moved to OBS, affecting the read and write performance of the system.

Possible Causes

The local disk capacity configured for cold-hot separation on the ClickHouseServer node is too small.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address of the hostname in **Location**.

Step 2 Expand the disk capacity of the node for which the alarm is generated.

Step 3 Check whether the alarm is cleared after disk expansion.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **ClickHouse** for the target cluster.

Step 6 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M personnel and provide the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.361 ALM-45475 A Single Replica Exists in the Kudu Data Table

Alarm Description

The system checks the replica status of the Kudu data table. This alarm is generated when a single replica is detected in the Kudu data table.

This alarm is cleared when all Kudu data tables have multiple replicas or no data.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45475	Major	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

A hardware fault, for example, a slow disk or a faulty disk, may cause Kudu data loss.

Handling Procedure

Step 1 Log in to the node where the MRS cluster client is installed as user **root** and switch to user **omm**.

su - omm

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Initialize the client.

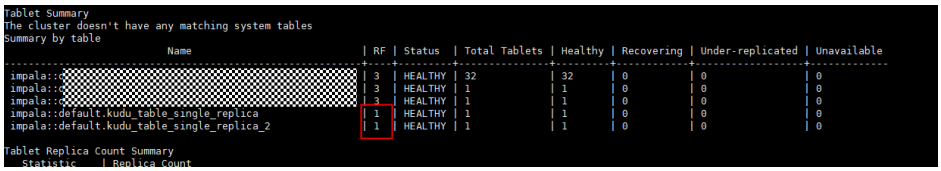
source /opt/Bigdata/client/bigdata_env

Step 3 If Kerberos authentication (in security mode) is enabled for the cluster, run the following command to authenticate the user. If Kerberos authentication (in normal mode) is not enabled for the cluster, skip this step.

kinit *{Username}*

Step 4 Run the following command to check whether there is a table whose **RF** is **1**:

kudu cluster ksck *<master_addresses>*



master_addresses: The value is in the format of *KuduMaster service IP address 1*:**7051,***KuduMaster service IP address 2*:**7051,***KuduMaster service IP address 3*:**7051**.

KuduMaster service IP address: You can log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kudu** > **Instances** to view the service IP address of the KuduMaster instance.

- If yes, consider whether to keep the single-replica table. You are advised to recreate the table and select at least three replicas for it.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.362 ALM-45476 Number of Tablets of the KuduTServer Process Exceeds the Threshold

Alarm Description

The system checks the Kudu monitoring metrics every 60 seconds. This alarm is generated when the number of tablets of the KuduTServer process exceeds the threshold.

This alarm is cleared when the number of tablets of the KuduTServer process becomes normal and the system considers that the Kudu service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45476	Minor	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the number of tablets exceeds the threshold, the query performance of the Kudu engine deteriorates.

Possible Causes

The TServer usage is too high or the TServer load is unbalanced.

Handling Procedure

Handle the Kudu instance exception.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm**, and check whether alarm **ALM-45476 Number of Tablets of the KuduTServer Process Exceeds the Threshold** exists.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > **Kudu**. Locate the threshold of this alarm and check whether the number of tablets of the KuduTServer process exceeds the threshold. If yes, change the threshold.

Step 3 Wait for several minutes, choose **O&M** > **Alarm**, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.363 ALM-45477 Failed to Restore Data After a Disk of Kudu Is Replaced

Alarm Description

This alarm is generated when Kudu fails to restore data by invoking the script in SetupTool after a disk is replaced.

This alarm is cleared when the data is successfully restored.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45477	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

Kudu fails to restore data, and historical data is unavailable.

Possible Causes

Kudu fails to restore the UUID or the data from the remote end.

Handling Procedure

Collect fault information.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 2 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 3 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

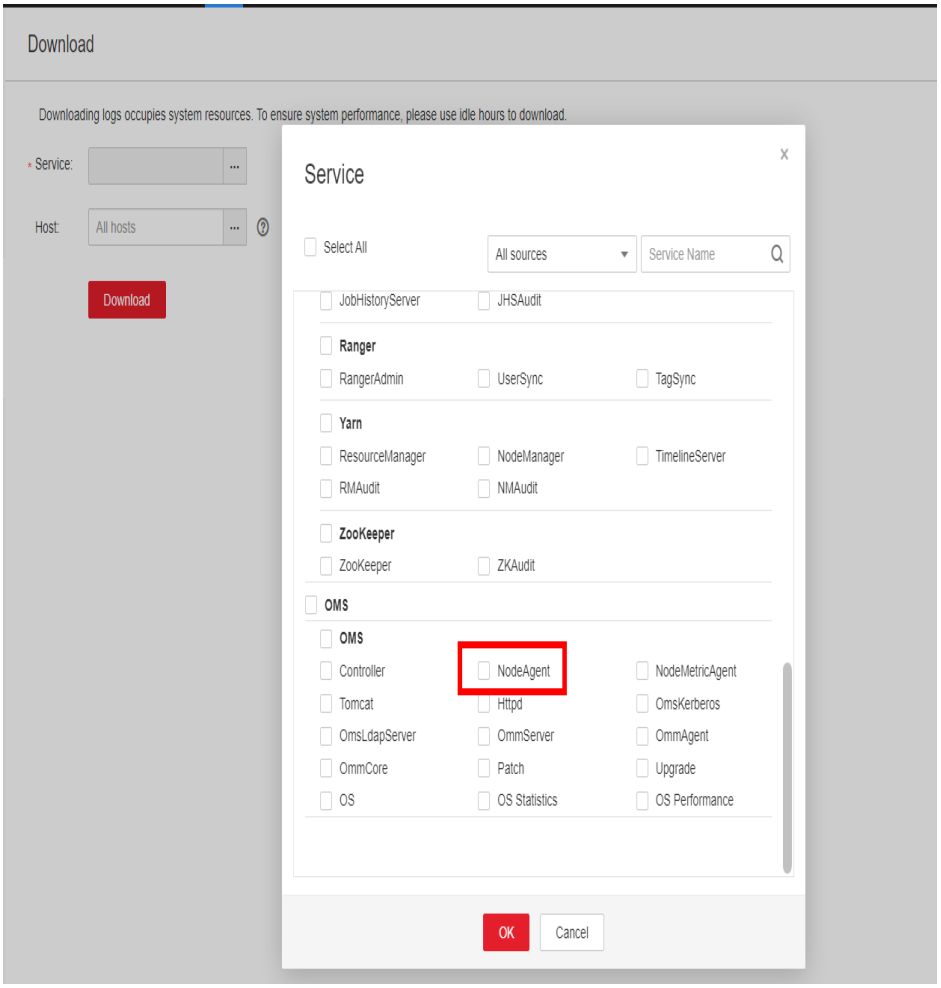
NOTE

Collect NodeAgent logs and check whether data is successfully restored after the disk replacement.

The logs are saved in the following directory:

/var/log/Bigdata/nodeagent/scriptlog/disk_manager_notify.log

Figure 9-13 Collecting NodeAgent logs



Step 4 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.364 ALM-45478 Kudu Failed Data Balancing

Alarm Description

The system periodically balances data among Kudu data tables. This alarm is generated when the system detects that the data balancing API returns a failure response.

This alarm is cleared when the Kudu data balancing API is successfully called.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45478	Major	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

Failed to call the data balancing API. The Kudu component may be abnormal.

Possible Causes

The segment or replica in the Kudu data table is lost.

Handling Procedure

Handle the Kudu instance exception.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. On the displayed page, locate alarm **ALM-45478 Kudu Failed Data Balancing**.

Step 2 In the **Location** field, record the host name and role name.

Step 3 Choose **Cluster** > **Services** > **Kudu** > **Instances**. Click the role name for the host name obtained in **Step 2** and restore the instance. Then, check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 On the Kudu web UI, check whether any segment or replica is missing.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.365 ALM-45479 Number of Tablets of the Tserver Process Exceeds the Threshold

Alarm Description

The system checks the Kudu service status every 60 seconds. This alarm is generated when the number of tablets of the Tserver process exceeds the threshold.

This alarm is cleared when the number of tablets of the Tserver process becomes normal and the system considers that the Kudu service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45479	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

If the number of tablets exceeds the threshold, the query performance of the Kudu engine deteriorates.

Possible Causes

The Tserver usage is too high or the Tserver load is unbalanced.

Handling Procedure

Handle the Kudu instance exception.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether alarm **ALM-45479 Number of Tablets of the Tserver Process Exceeds the Threshold** exists.

- If yes, go to **Step 2**.
- If no, go to **Step 6**.

Step 2 Choose **O&M** > **Alarm** > **Thresholds** > **Kudu**, locate the threshold of this alarm, compare the threshold with the monitored number of tablets in the Tserver process of the cluster, and handle the alarm accordingly.

- If the threshold is not properly set, change the threshold and go to **Step 5**.
- If the Tserver load is unbalanced, go to **Step 3**.
- If the Tserver usage is too high, delete obsolete tables or add Tserver nodes, and go to **Step 3**.

Step 3 Log in to the Kudu node where the number of tablets exceeds the threshold.

Step 4 Run the following commands to balance the load of the cluster (during off-peak hours, recommended):

su omm

cd /opt/Bigdata/FusionInsight_Kudu_xxx/install/FusionInsight-Kudu-xxx/kudu/bin

./kudu cluster rebalance <master_addresses> [-tables=<tables>]

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The preceding parameters can be obtained from the KuduMaster web UI. The **tables** parameter is optional.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.366 ALM-45480 Tablet Leaders of a Tserver Process Are Unevenly Distributed

Alarm Description

The system checks the Kudu service status every 60 seconds. This alarm is generated when the ratio of tablet leaders of a Tserver process to total tablet leaders in the cluster exceeds the threshold.

This alarm is cleared when the ratio of tablet leaders of a Tserver process to total tablet leaders in the cluster becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45480	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

If Tserver tablet leaders are unevenly distributed, the query performance of the Kudu engine deteriorates.

Possible Causes

- Tserver instances are added or restarted.
- The threshold is not properly set.

Handling Procedure

Handle the uneven distribution of tablet leaders.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether alarm **ALM-45480 Tablet Leaders of a Tserver Process Are Unevenly Distributed** exists.

- If yes, go to **Step 2**.
- If no, go to **Step 8**.

Step 2 Log in to the Kudu node where the number of tablet leaders exceeds the threshold.

Step 3 Run the following commands to manually adjust the number of tablet leaders:

su omm

cd /opt/Bigdata/FusionInsight_Kudu_*xxx*/install/FusionInsight-Kudu-xxx/kudu/bin

./kudu tablet leader_step_down *<master_addresses>* *<tablet_id>* **[-new_leader_uuid=***<new_tablet_server_uuid>***]**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- **master_addresses**: The value is in the format of **KuduMaster service IP address 1:7051,KuduMaster service IP address 2:7051,KuduMaster service IP address 3:7051**.

  KuduMaster service IP address: You can log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kudu** > **Instances** to view the service IP address of the KuduMaster instance.
- **tablet_id**: tablet ID to be adjusted.

  You can log in to Manager, choose **Cluster** > **Services** > **Kudu**, and click **KuduMaster(KuduMaster)** next to **KuduMaster WebUI** to access the KuduMaster web UI. Choose **Tables** on the menu bar, click the ID in the **Table Id** column, and obtain the target tablet ID in **Detail**.
- **new_tablet_server_uuid**: ID of the target tablet servers.

  You can log in to the KuduMaster web UI, choose **Tablet Servers** on the menu bar, and view the UUID.
- **new_leader_uuid**: You are advised to set the UUID of the node with a small number of tablet leaders.

Step 4 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Change the threshold.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and check whether alarm **ALM-45480 Tablet Leaders of a Tserver Process Are Unevenly Distributed** exists.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Choose **Cluster** > **Kudu** > **Configurations** > **All Configurations** > **KuduMaster**, locate the alarm threshold parameter **TABLET_LEADER_UNBALANCE_SCALE**, change the value, and rolling-restart all Kudu Masters to make the configuration take effect.

Step 7 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **Kudu** for the target cluster.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.367 ALM-45481 KuduTserver Has Full Disks

Alarm Description

The system checks Kudu disk metrics every 60 seconds. This alarm is generated when the number of fully occupied disks for a Tserver is not 0.

This alarm is automatically cleared when the number of fully occupied disks for the Tserver becomes 0.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45481	Major	Yes

Alarm Parameters

Type	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Users cannot use the Kudu service properly.

Possible Causes

The disk configuration cannot meet service requirements, and the disk space used by the Kudu service is insufficient.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the disk capacity and delete unnecessary files and data tables.

Step 1 Log in to FusionInsight Manager. In the alarm list, click the drop-down arrow in the row that contains the alarm, and view the role name and the IP address of the hostname in **Location**.

Step 2 Log in to the node for which the alarm is generated as user **root** and run the **df -h** command to check the mount directory where the disk usage is 100%.

Step 3 Run the **find / -xdev -size +500M;** command to view files larger than 500 MB in the directory obtained in **Step 2**. Check whether such files are mistakenly written to the directory.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Delete the mistakenly written files and check whether the alarm is cleared 2 minutes later.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Check whether Kudu contains unnecessary data tables.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Use **kudu table delete <master_addresses> <table_name> [-nomodify_external_catalogs]** to delete unnecessary tables.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 8 Contact the disk administrator to expand the disk capacity.

Step 9 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, select **Kudu** for the target cluster, and click **OK**.

Step 12 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.368 ALM-45585 IoTDB Service Unavailable

Alarm Description

The system checks the IoTDB service status every 300 seconds. This alarm is generated when the IoTDB service is unavailable. This alarm is cleared when the IoTDB service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45585	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The IoTDB service is unavailable for read and write operations.

Possible Causes

- The KrbServer service is abnormal.
- More than 50% of IoTDBServer instances are faulty.

Handling Procedure

Check whether the KrbServer service on which the IoTDB depends is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether ALM-25500 KrbServer Service Unavailable exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by referring to **ALM-25500 KrbServer Service Unavailable**.

Step 4 After ALM-25500 is cleared, wait a few minutes and check whether the alarm HetuServer Service Unavailable is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether IoTDBServer instances are faulty.

Step 5 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Instance**.

Step 6 Check whether the percentage of faulty IoTDBServer instances exceeds 50%. If yes, restart the faulty IoTDBServer instances and check whether the status is restored.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.369 ALM-45586 IoTDBServer Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the IoTDBServer process status every 60 seconds. The alarm is generated when the heap memory usage of the IoTDBServer process exceeds the threshold (90% of the maximum memory).

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45586	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The data read and write performance of the IoTDBServer process deteriorates, memory overflow occurs, and the IoTDBServer process breaks down.

Possible Causes

The heap memory of the IoTDBServer process is overused or the heap memory is inappropriately allocated.

Handling Procedure

Check the heap memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **45586**, view the role name in **Location**, and check the instance IP address.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Instance**. Click the IoTDBServer for which the alarm is generated to go to **Dashboard**. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **Memory**. In the dialog box that is displayed, select **IoTDBServer Heap Memory Resource Percentage**, and click **OK**. Check whether the used non-heap memory of the IoTDBServer process reaches 90% (by default) of the maximum non-heap memory specified for IoTDBServer.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Configuration**, click **All Configurations**, choose **IoTDBServer** > **System**, and increase the value of **-Xmx** in the **GC_OPTS** parameter.

NOTE

- The default value of -**Xmx** is **2G**.
- If this alarm is occasionally generated, increase the value by 0.5 times. If this alarm is frequently generated, double the value.
- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.370 ALM-45587 IoTDBServer GC Duration Exceeds the Threshold

Alarm Description

The system checks the GC duration of the IoTDBServer process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default) for three consecutive times. You can choose **O&M** > **Alarm** > **Threshold Configuration** > *Name of the desired cluster* > **IoTDB** > **GC** > **Total GC duration of IoTDBServer process (IoTDBServer)** to change the threshold. This alarm is cleared when the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45587	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The IoTDBServer process may be unavailable for data read and write operations.

Possible Causes

The heap memory of the IoTDBServer process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC duration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **45587**, view the role name in **Location**, and check the instance IP address.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Instance**. Click the IoTDBServer for which the alarm is generated to go to **Dashboard**. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **GC**. In the dialog box that is displayed, select **Garbage Collection (GC) Time of IoTDBServer**, and click **OK**. Check whether the GC time of the IoTDBServer process is greater than 12 seconds.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Configuration**, click **All Configurations**, choose **IoTDBServer** > **System**, and increase the value of **-Xmx** in the **GC_OPTS** parameter.

NOTE

- The default value of **-Xmx** is **2G**.
- If this alarm is occasionally generated, increase the value by 0.5 times. If this alarm is frequently generated, double the value.
- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect the fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.371 ALM-45588 IoTDBServer Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the IoTDBServer service every 60 seconds. This alarm is generated when the direct memory usage of the IoTDBServer instance exceeds the threshold (90% of the maximum memory) for five consecutive times. This alarm is cleared when the IoTDBServer direct memory usage is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45588	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Direct memory overflow may cause service unavailability, and the IoTDBServer process may be unavailable for data read and write operations.

Possible Causes

The direct memory of the IoTDBServer process is overused or the direct memory is inappropriately allocated.

Handling Procedure

Check the direct memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to FusionInsight Manager and choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**. On the page that is displayed, locate the row containing the alarm whose **Alarm ID** is **45588**, view the role name in **Location**, and check the instance IP address.

Step 2 Choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Instance**. Click the IoTDBServer for which the alarm is generated to go to **Dashboard**. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **Memory**. In the dialog box that is displayed, select **IoTDBServer Direct Buffer Resource Percentage**, and click **OK**.

Step 3 Check whether the direct memory used by the IoTDBServer reaches the threshold (90% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **IoTDB** > **Configuration**, click **All Configurations**, choose **IoTDBServer** > **System**, increase the value of **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter as required, and save the configuration.

NOTE

- If this alarm is generated, the direct memory configured for the IoTDBServer process cannot meet the requirements of the IoTDBServer process.
- You are advised to set **-XX:MaxDirectMemorySize** in **GC_OPTS** to twice the direct memory used by the IoTDBServer process. (You can change the value based on the actual service scenario.)
- To obtain the size of the direct memory used by the IoTDBServer process, choose **Customize** > **Memory** > **IoTDBServer Direct Memory Resource Status**. If **GC_OPTS** does not contain the **-XX:MaxDirectMemorySize** parameter, add it manually.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect the fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **IoTDBServer** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.372 ALM-45589 ConfigNode Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the ConfigNode process every 60 seconds. This alarm is generated when the heap memory usage of the ConfigNode process exceeds the threshold (90% of the maximum memory). This alarm is cleared when the heap memory usage of the ConfigNode process is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45589	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The data read and write performance of the ConfigNode process deteriorates, memory overflow occurs, and the ConfigNode process breaks down.

Possible Causes

The heap memory configured for the node is improper. As a result, the usage exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check the heap memory configuration.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click  in front of this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **Cluster** > **Services** > **IoTDB**. Click **Instance**, click the ConfigNode corresponding to the IP address obtained in **Step 1**, and check whether **ConfigNode Heap Memory Usage** on the **Dashboard** tab page reaches the threshold specified for the ConfigNode process.

If the chart is not displayed, click the drop-down list in the upper right corner of the chart area and choose **Customize** > **Memory**. In the dialog box that is displayed, select **ConfigNode Heap Memory Usage** and click **OK**.

NOTE

You can choose **O&M** > **Alarm** > **Threshold Configuration** > *Name of the desired cluster* > **IoTDB** > **Memory** > **ConfigNode Heap Memory Usage (ConfigNode)** to view the threshold.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Choose **Cluster** > **Services** > **IoTDB**. Click **Configurations** then **All Configurations**, click **ConfigNode**, and choose **System**. Set **-Xmx** in **GC_OPTS** to a larger value and save the configuration.

NOTE

- The default value of -**Xmx** is **2G**.
- If this alarm is occasionally generated, increase the value of **-Xmx** by 0.5 times. If this alarm is frequently generated, double the value of **-Xmx**.
- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Click **Dashboard**. Click **Restart Service** to restart the IoTDB service for the configuration to take effect.

NOTICE

During the restart of the IoTDB service, read and write requests are interrupted.

Step 5 Wait for about 120 seconds and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.373 ALM-45590 ConfigNode GC Duration Exceeds the Threshold

Alarm Description

The system checks the GC duration of the ConfigNode process every 60 seconds. This alarm is generated when the GC duration exceeds the threshold (12 seconds by default) for three consecutive times. This alarm is cleared when the GC duration is less than the threshold.

NOTE

You can choose **O&M** > **Alarm** > **Threshold Configuration** > *Name of the desired cluster* > **IoTDB** > **GC** > **Total GC duration of ConfigNode process (ConfigNode)** to increase the threshold by 20% each time.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45590	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The data read and write performance of the ConfigNode process may deteriorate.

Possible Causes

The heap memory configured on the node is improper. As a result, GC occurs frequently.

Handling Procedure

Check the heap memory configuration.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click in the row containing this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **Cluster** > **Services** > **IoTDB**. Click **Instance**, click the ConfigNode corresponding to the IP address obtained by **Step 1**. Switch to the **Dashboard** tab page, locate the **Total GC Duration of ConfigNode** chart, and check whether the GC duration of the ConfigNode process exceeds the threshold.

If the GC duration of ConfigNode is not displayed, click the drop-down list in the upper right corner of the chart area and choose **Customize** > **GC**. In the displayed dialog box, select **Total GC Duration of ConfigNode** and click **OK**.

NOTE

You can choose **O&M** > **Alarm** > **Threshold Configuration** > *Name of the desired cluster* > **IoTDB** > **GC** > **Total GC duration of ConfigNode process (ConfigNode)** to view the threshold.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Choose **Cluster** > **Services** > **IoTDB**. Click **Configurations** then **All Configurations**, click **ConfigNode**, and choose **System**. Set **-Xmx** in **GC_OPTS** to a larger value and save the configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- The default value of -**Xmx** is **2G**.
- If this alarm is occasionally generated, increase the value by 0.5 times. If this alarm is frequently generated, double the value.
- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Click **Dashboard**. Click **Restart Service** to restart the IoTDB service for the configuration to take effect.

NOTICE

During the restart of the IoTDB service, read and write requests are interrupted.

Step 5 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 8 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.374 ALM-45591 ConfigNode Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the ConfigNode process every 60 seconds. This alarm is generated when the direct memory usage of the ConfigNode exceeds the threshold for five consecutive times. That is, the direct

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

memory configured for ConfigNode cannot meet service requirements. This alarm is cleared when the direct memory usage of ConfigNode is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45591	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Direct memory overflow may cause instance unavailability, and the ConfigNode process may be unavailable for data read and write operations.

Possible Causes

The direct memory configured for the node is improper. As a result, the usage exceeds the threshold.

Handling Procedure

Check the direct memory configuration.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click in front of this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **Cluster** > **Services** > **IoTDB**. Click **Instance**, click the ConfigNode corresponding to the IP address obtained in **Step 1**, and check whether **ConfigNode Direct Memory Usage** on the **Dashboard** tab page reaches the

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

threshold specified for the ConfigNode process (90% of the maximum direct memory by default).

If the chart is not displayed, click the drop-down list in the upper right corner of the chart area and choose **Customize** > **Memory**. In the dialog box that is displayed, select **ConfigNode Direct Memory Usage** and click **OK**.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Services** > **IoTDB**. Click **Configurations** then **All Configurations**. Click **ConfigNode** and select **System**. Set **-XX:MaxDirectMemorySize** in **GC_OPTS** to a larger value as required and save the configuration.

NOTE

- You are advised to set **-XX:MaxDirectMemorySize** in **GC_OPTS** to twice the direct memory used by the ConfigNode process. (You can change the value based on the actual service scenario.)
- To obtain the size of the direct memory used by the ConfigNode process, choose **Customize** > **Memory** > **ConfigNode Direct Memory Resource Status**.
- If **GC_OPTS** does not contain the **-XX:MaxDirectMemorySize** parameter, add it.

Step 4 Restart the affected IoTDB service or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

NOTICE

During the restart of the IoTDB service or instance, read and write requests are interrupted.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **ConfigNode** for the destination cluster.

Step 7 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.375 ALM-45592 IoTDBServer RPC Execution Duration Exceeds the Threshold

Alarm Description

The system checks the RPC execution duration of the IoTDBServer process every 60 seconds. This alarm is generated when the execution duration exceeds the threshold. This alarm is cleared when the RPC execution time of the IoTDBServer process is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45592	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The data read and write performance of the IoTDBServer process deteriorates.

Possible Causes

The processing duration of an IoTDBServer RPC request exceeds the threshold. Logs need to be further analyzed to locate the cause.

Handling Procedure

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2023

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click in front of this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **O&M** > **Log** > **Download**.

Step 3 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 4 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the host queried in **Step 1**, and click **OK**.

Step 5 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 6 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.376 ALM-45593 IoTDBServer Flush Execution Duration Exceeds the Threshold

Alarm Description

This alarm is generated when the data flush duration exceeds the threshold. This alarm is cleared when the flush duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45593	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Data writes are blocked, and write performance deteriorates.

Possible Causes

The IoTDB flushing on the node is slow. You need to further analyze logs.

Handling Procedure

Collect fault information.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click in front of this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **O&M** > **Log** > **Download**.

Step 3 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 4 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the host queried in **Step 1**, and click **OK**.

Step 5 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 6 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2025

9.11.377 ALM-45594 IoTDBServer Intra-Space Merge Duration Exceeds the Threshold

Alarm Description

This alarm is generated when the merge duration in the space exceeds the threshold. This alarm is cleared when the merge duration in the space is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45594	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Data writes are blocked, and write performance deteriorates.

Possible Causes

The merge task in the IoTDB space of the node is slow. You need to further analyze logs.

Handling Procedure

Collect fault information.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click in front of this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **O&M** > **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 4 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the host queried in **Step 1**, and click **OK**.

Step 5 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 6 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.378 ALM-45595 IoTDBServer Cross-Space Merge Duration Exceeds the Threshold

Alarm Description

This alarm is generated when the cross-space merge duration exceeds the threshold. This alarm is cleared when the cross-space merge duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45595	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Data writes are blocked, which causes write performance deterioration.

Possible Causes

The IoTDB cross-space merge task on the node is slow. You need to further analyze logs.

Handling Procedure

Collect fault information.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the real-time alarm list, click  in front of this alarm and view the role name and instance IP address in **Location**.

Step 2 Choose **O&M** > **Log** > **Download**.

Step 3 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 4 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the host queried in **Step 1**, and click **OK**.

Step 5 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 6 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.379 ALM-45596 Procedure Execution Failed

Alarm Description

Procedures are the tasks managed and executed by the ConfigNode leader. This alarm is generated when a procedure fails to be executed. This alarm is cleared when the procedure is successfully executed.

NOTE

This section applies to MRS 3.3.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45596	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
ProcedureInformation	Specifies the procedure-related information.

Impact on the System

Operation consistency is affected.

Possible Causes

- The task for adding IoTDB replicas fails to be executed.
- The task for deleting the storage group fails to be executed.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, locate this alarm, and click .

Step 2 Check the value of **ProcedureInformation** in **Location**. The value starts with the procedure type and contains main information about the procedure.

Check whether the task for adding replicas fails.

Step 3 Check whether the value of **ProcedureInformation** starts with **AddRegionProcedure** or **ReJoinDataNodeProcedure**.

- If yes, the task fails. Go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Wait for half an hour. If the region is successfully added, the alarm is automatically cleared. Otherwise, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the task for deleting the storage group fails.

Step 5 Check whether the value of **ProcedureInformation** starts with **DeleteStorageGroupProcedure**.

- If yes, the storage group fails to be deleted. Go to **Step 6**.
- If no, go to **Step 7**.

Step 6 Delete the storage group displayed in **ProcedureInformation** again on the IoTDB client. If the deletion is successful, the alarm is automatically cleared. Otherwise, go to **Step 7**.

Collect fault information.

Step 7 Choose **Cluster** > **Services** > **IoTDB** > **Instance** to view the hosts where all IoTDBServer and ConfigNode instances are located.

Step 8 Choose **O&M** > **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, select **IoTDB** for the target cluster, and click **OK**.

Step 10 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the host queried in **Step 7**, and click **OK**.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.380 ALM-45615 CDL Service Unavailable

Alarm Description

The system checks the CDL health status every 60 seconds. This alarm is generated when the CDL health status is **DOWN**. This alarm is cleared when the system detects that the CDL health status is **UP**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45615	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The CDL service is abnormal. You cannot use FusionInsight Manager to perform cluster operations on the CDL service. The CDL service function is unavailable.

Possible Causes

All CDLService or CDLConnector instances of the CDL service are abnormal, and the Kafka service is unavailable.

Handling Procedure

Check whether the Kafka service on which the CDL service depends is abnormal.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether ALM-38000 Kafka Service Unavailable exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by referring to **ALM-38000 Kafka Service Unavailable**.

Step 4 After the alarm is cleared, wait a few minutes and check whether the alarm HetuServer Service Unavailable is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether CDL instances are faulty.

Step 5 On FusionInsight Manager, choose **Cluster** > *Name of the desired cluster* > **Service** > **CDL** > **Instance**.

Step 6 Check whether all CDLService and CDLConnector instances are faulty.

- If yes, restart the CDL service and choose **Cluster** > *Name of the desired cluster* > **Services** > **CDL** > **More** > **Restart Service**. If the fault persists after the restart, go to **Step 7** and contact O&M personnel to check CDL logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **CDL** for the target cluster.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

After the service is restored, the system automatically clears the alarm.

Related Information

None

9.11.381 ALM-45616 CDL Job Execution Exception

Alarm Description

The system checks whether a CDL job is normal every 60 seconds. This alarm is reported when the CDL job is abnormal. This alarm is cleared when the job is restored or stopped.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45616	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
JobName	Specifies the job for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Username	Specifies the username of the job for which the alarm was generated.

Impact on the System

CDL jobs fail, and real-time data integration is interrupted.

Possible Causes

The CDL task fails to be executed due to incorrect parameter settings or other reasons. On the **Job Management** page of the CDL web UI, locate the row where the job is located and click **Failed**/**Abnormal running** in the **Status** column to view the failure cause, or view the failure cause in the logs.

Handling Procedure

Step 1 Log in to FusionInsight Manager as a user who has the CDL job creation or administrator permission.

Step 2 Choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**, click ⌄ in the row where **Alarm ID** is **45616**, and view the name of the job for which this alarm is generated in **Location**.

Step 3 Choose **Cluster** > **Services** > **CDL** and click the link next to **CDLService UI** to go to the CDL web UI.

Step 4 Locate the row where the failed job is located based on the job name obtained in **Step 2**, and click **Abnormal running** or **Failed** in the **Status** column.



Step 5 On the page that is displayed, view the error information and rectify the fault. For example, **Figure 9-14** shows that the task running on Yarn is manually killed. For details, see trace error information, as shown in **Figure 9-15**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-14 CDL job exception

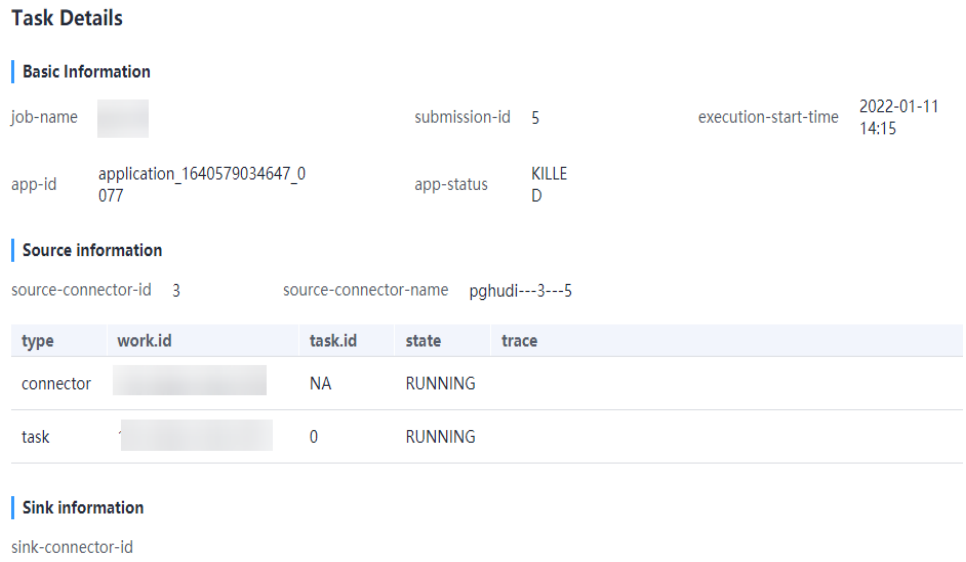
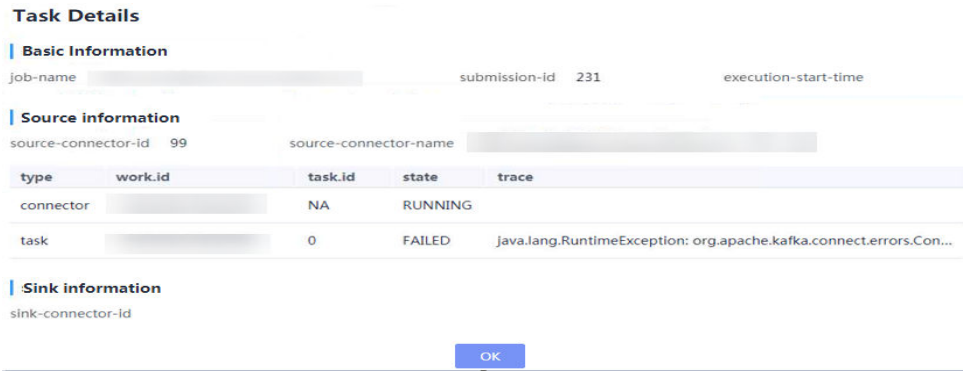


Figure 9-15 Trace error information



Step 6 Rectify the fault based on the error information, execute the task again, and check whether the task can be executed successfully.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect the fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Select **CDL** in the required cluster for **Service**.

Step 9 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

After the job is successfully restored or stopped, the alarm is cleared if it has been reported.

Related Information

None

9.11.382 ALM-45617 Data Queued in the CDL Replication Slot Exceeds the Threshold

Alarm Description

If too many WALs are stacked in PostgreSQL, the PostgreSQL disk space may be used up. The system checks whether the amount of data queued in the replication slot configured for a CDL job exceeds the threshold every 5 minutes. This alarm is generated when the amount of data queued in the replication slot exceeds the threshold. This alarm is cleared when the number of data queued in the replication slot falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45617	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
JobName	Specifies the job for which the alarm was generated.
DBName	Specifies the database for which the alarm was generated.
SlotName	Specifies the database replication slot for which the alarm was generated.
Lag	Specifies the data queued in the slot.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The disk space of the source PostgreSQL database may be used up and the database cannot provide services.

Possible Causes

The CDL job is abnormal, and data processing stops; the source database is updated quickly, and CDL data processing is slow.

Handling Procedure

Step 1 Log in to FusionInsight Manager as a user who has the CDL job creation or administrator permission.

Step 2 Choose **O&M**. In the navigation pane on the left, choose **Alarm** > **Alarms**, click  in the row where **Alarm ID** is **45617**, and view the name of the job for which this alarm is generated in **Location**.

Step 3 Check whether **ALM-45616 CDL Job Execution Exception** is displayed in the alarm list.

- If yes, handle the alarm by referring to **ALM-45616 CDL Job Execution Exception**.
- If no, go to **Step 4**.

Step 4 Choose **Cluster** > **Services** > **CDL**. Click the link next to **CDLService UI** to go to the CDL web UI and check whether the job is displayed in the job list based on its name obtained in **Step 2**.

- If yes, check whether the job is abnormal.
  - If it is abnormal, go to **Step 5**.
  - If it is not, data processing is slow. Contact O&M personnel.
- If no, go to **Step 7**.

Step 5 Click **Abnormal** or **Failed** in the row where the job is located and rectify the fault based on the error information displayed on the page.

Step 6 After rectifying the fault, run the job again and check whether the job can be executed successfully.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **CDL** for the target cluster.

Step 9 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is cleared when the amount of data queued in the replication slot is less than the threshold. You do not need to manually clear the alarm.

Related Information

None

9.11.383 ALM-45635 FlinkServer Job Execution Failure

This section applies to MRS 3.1.2 or later.

Alarm Description

The system checks whether FlinkServer jobs fail to be executed every 10 seconds. This alarm is generated when a FlinkServer job fails. This alarm is cleared when the job is successfully restarted.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45635	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
JobName	Specifies the job for which the alarm was generated.

Impact on the System

This alarm is a job-level alarm and does not affect FlinkServer. You need to view Flink job logs to find out the failure cause.

Possible Causes

You can view failure causes in specific logs.

Handling Procedure

Step 1 Log in to Manager as a user who has the FlinkServer management permission.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the Yarn page.

Step 3 Locate the failed job based on its name displayed in **Location**, search for and record the application ID of the failed job, and check whether the job logs are available on the Yarn page.

If yes, go to **Step 4**.

If no, go to **Step 6**.

Step 4 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

NOTE

You can also log in to Manager as a user who has the FlinkServer management permission, choose **Cluster** > **Services** > **Flink**, click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management** and choose **More** > **Job Monitoring** in the **Operation** column to view the TaskManager logs.

Step 5 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, select **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***User name*/**logs/***Application ID of the failed job* directory.

Step 7 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

After the job is successfully restarted, the alarm is cleared if it has been reported.

Related Information

None

9.11.384 ALM-45636 Flink Job Checkpoints Keep Failing

This section applies to MRS 3.1.2 or a version between 3.1.2 and 3.3.0.

Alarm Description

The system checks the number of consecutive checkpoint failures based on the configured alarm checking interval. This alarm is generated when the number of consecutive checkpoint failures of a FlinkServer job reaches the configured threshold. This alarm is cleared when checkpoints are recovered or the job is successfully restarted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45636	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.
Username	Specifies the username of the job for which the alarm is generated.

Impact on the System

The Flink job may fail. You need to check the status and logs of the Flink job to locate the fault. This is a job-level alarm and has no impact on FlinkServer.

Possible Causes

You can view failure causes in specific logs.

Handling Procedure

Step 1 Log in to Manager as a user who has the FlinkServer management permission.

Step 2 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the Yarn page.

Step 3 Locate the failed job based on its name displayed in **Location**, search for and record the application ID of the failed job, and check whether the job logs are available on the Yarn page.

If yes, go to **Step 4**.

If no, go to **Step 6**.

Step 4 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

You can also log in to Manager as a user who has the FlinkServer management permission, choose **Cluster** > **Services** > **Flink**, click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management** and choose **More** > **Job Monitoring** in the **Operation** column to view the TaskManager logs.

Step 5 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, select **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***User name*/**logs/***Application ID of the failed job* directory.

Step 7 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

This alarm is cleared when Flink job checkpoints are recovered or the job is successfully restarted.

Related Information

None

9.11.385 ALM-45636 Number of Consecutive Checkpoint Failures of a Flink Job Exceeds the Threshold

This section applies to MRS 3.3.1 or later.

Alarm Description

The system checks the number of consecutive checkpoint failures based on the configured alarm checking interval. This alarm is generated when the number of consecutive checkpoint failures of a FlinkServer job reaches the configured threshold. This alarm is cleared when checkpoints are recovered or the job is successfully restarted.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45636	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	ApplicationName	Specifies the name of the application for which the alarm was generated.
	JobName	Specifies the job for which the alarm was generated.
	UserName	Specifies the username for which the alarm was generated.
Additional Information	ThresholdValue	Specifies the threshold value for triggering the alarm.
	CurrentValue	Specifies the value that triggered the alarm.

Impact on the System

The Flink job may fail. You need to check the status and logs of the Flink job to locate the fault. This is a job-level alarm and has no impact on FlinkServer.

Possible Causes

You can view failure causes in specific logs.

Handling Procedure

Step 1 Log in to Manager as a user who has the FlinkServer management permission.

Step 2 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 3 Locate the failed task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the native Yarn page.

Figure 9-16 Application ID of a job



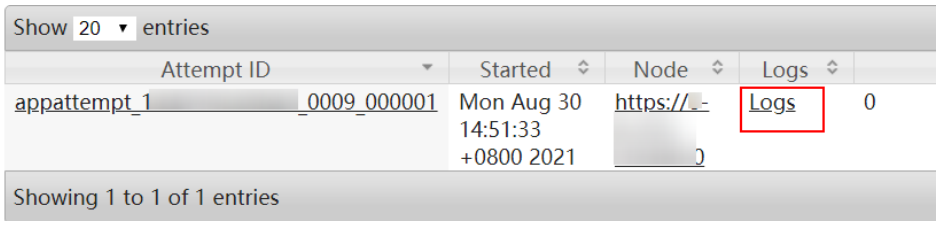
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

   Figure 9-17 Clicking Logs

2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

   Figure 9-18 Clicking the ID in the Attempt ID column

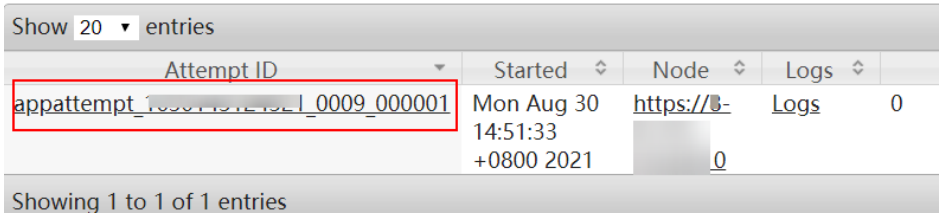
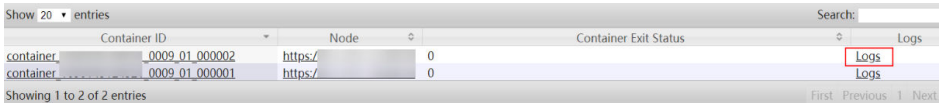
   

   Figure 9-19 Clicking Logs

   NOTE

   You can also log in to Manager as a user who has the FlinkServer management permission. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

Step 5 View the logs of the failed job to rectify the fault, or contact the O&M engineers and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**logs/***Application ID of the failed job* directory.

Step 7 View the logs of the failed job to rectify the fault, or contact the O&M engineers and send the collected fault logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is cleared when FlinkServer job checkpoints are recovered or the job is successfully restarted.

Related Information

None.

9.11.386 ALM-45637 FlinkServer Task Is Continuously Under Back Pressure

This section applies to MRS 3.1.2 or later.

Alarm Description

The system checks the back pressure duration of FlinkServer tasks based on the configured alarm checking interval. This alarm is generated when the back pressure duration of a FlinkServer task reaches the configured threshold. This alarm is cleared when the task back pressure is recovered or the job is successfully restarted.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45637	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
JobName	Specifies the job for which the alarm was generated.

Impact on the System

Continuous back pressure of Flink jobs may cause performance problems or checkpoint failures. As a result, Flink jobs fail. You need to check the status and logs of the Flink jobs to locate the cause. This is a job-level alarm and has no impact on FlinkServer.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

You can view the causes in the specific logs.

Handling Procedure

Step 1 Log in to Manager as a user who has the FlinkServer management permission.

Step 2 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the Yarn page.

Step 3 Locate the failed job based on its name displayed in **Location**, search for and record the application ID of the failed job, and check whether the job logs are available on the Yarn page.

If yes, go to **Step 4**.

If no, go to **Step 6**.

Step 4 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

   NOTE

   You can also log in to Manager as a user who has the FlinkServer management permission, choose **Cluster** > **Services** > **Flink**, click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management** and choose **More** > **Job Monitoring** in the **Operation** column to view the TaskManager logs.

Step 5 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, select **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***User name*/**logs/***Application ID of the failed job* directory.

Step 7 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

This alarm is cleared when FlinkServer task back pressure is recovered or the job is successfully restarted.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.387 ALM-45638 Number of Restarts After FlinkServer Job Failures Exceeds the Threshold

Alarm Description

The system checks the number of FlinkServer job restarts based on the alarm checking interval. This alarm is generated when the number exceeds the configured threshold. This alarm is cleared when the job is restarted.

This alarm handling guide is targeted for MRS clusters of versions 3.1.2 to 3.2.0. For those MRS clusters of version 3.2.0 or later, see ALM-45638 Number of Restarts After Flink Job Failures Exceeds the Threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45638	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
JobName	Specifies the job for which the alarm was generated.

Impact on the System

Flink jobs are frequently restarted due to the failures. You need to locate the cause. This is a job-level alarm and has no impact on FlinkServer.

Possible Causes

You can view the causes in the specific logs.

Handling Procedure

Step 1 Log in to Manager as a user who has the FlinkServer management permission.

Step 2 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the Yarn page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Locate the failed job based on its name displayed in **Location**, search for and record the application ID of the failed job, and check whether the job logs are available on the Yarn page.

If yes, go to **Step 4**.

If no, go to **Step 6**.

Step 4 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

NOTE

You can also log in to Manager as a user who has the FlinkServer management permission, choose **Cluster** > **Services** > **Flink**, click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management** and choose **More** > **Job Monitoring** in the **Operation** column to view the TaskManager logs.

Step 5 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, select **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***User name*/**logs/***Application ID of the failed job* directory.

Step 7 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

After the fault is rectified, the system does not automatically clear this alarm, and you need to manually clear the alarm.

Related Information

None

9.11.388 ALM-45638 Number of Restarts After Flink Job Failures Exceeds the Threshold

Description

The system checks the number of FlinkServer job restarts at intervals defined by the parameter **metrics.reporter.alarm.interval** (default value: 30s). This alarm is generated when the number exceeds the configured threshold. This alarm is cleared when the job is restarted.

The alarm threshold is calculated as follows: Threshold = Percentage of failed restarts that trigger an alarm x Maximum number of allowed job restarts.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The default threshold in MRS 3.2.1 or later is 3 (calculated by rounding up 80% x 3).

- Percentage of failed restarts that trigger an alarm: specified by the parameter **metrics.reporter.alarm.job.alarm.failure.restart.rate**. The default value is 80.
- Maximum number of allowed job restarts: The value varies depending on the restart strategy configured for the current Flink job. The default restart strategy is **fixed-delay** in MRS 3.2.1 or later.
  - If the restart strategy is **none**, this alarm is not triggered. (The default restart strategy is **none** in versions earlier than MRS 3.2.1).
  - If the restart strategy is **failure-rate**, the default maximum number of allowed restarts is **1** (**restart-strategy.failure-rate.max-failures-per-interval**).

    In this case, the alarm threshold is 1 (calculated by rounding up 80% × 1). This means that if the job restarts even once, the system will generate an alarm.
  - If the restart strategy is **fixed-delay**, the default maximum number of allowed restarts is **3** (**restart-strategy.fixed-delay.attempts**).

    In this case, the alarm threshold is 3 (calculated by rounding up 80% × 3). This means that if the job restarts three times, the system will generate an alarm.

This section applies to MRS 3.2.0-LTS.1 or later.

Attribute

Alarm ID	Alarm Severity	Auto Clear
45638	Major	Yes

Parameters

Name	Meaning
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.
Username	Specifies the username of the job for which the alarm is generated.

Impact on the System

If the number of Flink job restarts exceeds the configured threshold, the job fails repeatedly and restarts frequently. In such cases, the job developer should check

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the job logs to identify the root cause of the failure. This alarm is a Flink job alarm and does not affect FlinkServer.

Possible Causes

You can view the causes in the specific logs.

Procedure

Step 1 Log in to the MRS Manager using a user who has the FlinkServer management operation permission, Manager interface management permission, and Yarn WebUI access permission.

Step 2 Open the FusionInsight Manager page. In the alarm list, click ⌄ in the row where this alarm is located to view the task name in the alarm details.

Step 3 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the Yarn page.

Step 4 Locate the failed job based on its name displayed in **Location**, search for and record the application ID of the failed job, and check whether the job logs are available on the Yarn page.

If yes, go to **Step 5**.

If no, go to **Step 6**.

Step 5 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

   NOTE

   You can also log in to Manager as a user who has the FlinkServer management permission, choose **Cluster** > **Services** > **Flink**, click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management** and choose **More** > **Job Monitoring** in the **Operation** column to view the TaskManager logs.

Step 6 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 7 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page.

Step 8 Select **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***User name*/**logs/***Application ID of the failed job* directory.

Step 9 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearing

After the fault is rectified, the system does not automatically clear this alarm, and you need to manually clear the alarm.

Related Information

None

9.11.389 ALM-45639 Checkpointing of a Flink Job Times Out

Alarm Description

The system checks the checkpointing timeout of Flink jobs every 30 seconds. This alarm is generated if the checkpointing timeout of a Flink job is longer than the threshold (600 seconds by default). This alarm is cleared when the checkpointing timeout of a job is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45639	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName (available in MRS 3.2.1 or later)	Specifies the name of the application for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.
UserName	Specifies the username for which the alarm is generated.

Impact on the System

The checkpointing fails. You need to locate the cause. This is a job-level alarm and has no impact on FlinkServer.

Possible Causes

The job may be in the sub-healthy state. The possible causes are as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The memory for the TaskManager of the job is insufficient.
- The state memory is too large, making checkpointing time-consuming.

Handling Procedure

Step 1 Log in to Manager as a user who has the FlinkServer management permission.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45639 Checkpointing of a Flink Job Times Out**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 4 Locate the failed task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view TaskManager logs.

   NOTE

   You can also log in to Manager as a user who has the FlinkServer management permission. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

Step 6 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 7 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**logs/***Application ID of the failed job* directory.

Step 8 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

This alarm is cleared when the checkpointing timeout a Flink job is less than or equal to the threshold.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.390 ALM-45640 FlinkServer Heartbeat Interruption Between the Active and Standby Nodes

This section applies to MRS 3.2.0 or later.

Alarm Description

This alarm is generated when the FlinkServer active node or standby node does not receive heartbeat messages from the peer for 30 seconds (heartbeat interruption duration configured in keepalive).

This alarm is cleared when the heartbeat recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45640	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The impact varies depending on the cause. If the heartbeat is interrupted due to active node faults, the standby node will take over services as the active node. If the heartbeat is interrupted due to other reasons, for example, network problems, two active nodes may exist because the standby node has become the active node. Data synchronization between the active and standby nodes is abnormal, but FlinkServer can still provide services.

Possible Causes

- The active or standby FlinkServer instance is in the stopped state.
- The NIC of the floating IP address of the HA system used by the FlinkServer node is incorrectly configured. FlinkServer fails to be started.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The link between the active and standby FlinkServer nodes is abnormal.

Handling Procedure

Check the status of the active and standby FlinkServer instances.

Step 1 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **Flink** > **Instance**, and check the state of FlinkServer is normal.

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Select the abnormal FlinkServer instance and start the instance. After the instance is started, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

NOTICE

During the restart, the FlinkServer instance cannot provide services, but submitted jobs are not affected.

Check whether the link between the standby FlinkServer nodes is normal.

Step 3 Choose **Cluster** > **Services** > **Flink** > **Instance**, and check the two service IP addresses of FlinkServer.

Step 4 Log in to the server where the abnormal FlinkServer instance locates as the **root** user.

Step 5 Run the following command to check whether the server of the other FlinkServer instance is reachable:

ping *IP address of the other FlinkServer instance*

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 6 Ask the network administrator to handle the network exception.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the logs of the node where the abnormal FlinkServer instance locates contains error information.

Step 8 Log in to the server where the abnormal FlinkServer instance locates as the **root** user.

Step 9 Open the log file in the default directory **/var/log/Bigdata/flink/flinkserver/prestart.log** and check whether there is error message **Float ip** *x.x.x.x* **is invalid**.

- If yes, go to **Step 10**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 12**.

Step 10 On FusionInsight Manager, choose **Cluster** > **Services** > **Flink** > **Configurations** > **All Configurations** and search for **flink.ha.floatip**. Change the parameter value to the correct floating IP address, save the configuration, and restart the Flink service.

NOTICE

- Contact the network engineer to obtain the new floating IP address.
- During the service restart, FlinkServer cannot provide services, but submitted jobs are not affected.
- During the restart, the FlinkServer instance cannot provide services, but submitted jobs are not affected.

Step 11 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect the fault information.

Step 12 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 13 Select the Flink service in the required cluster for **Service**.

Step 14 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 15 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 16 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.391 ALM-45641 Data Synchronization Exception Between the Active and Standby FlinkServer Nodes

This section applies to MRS 3.2.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Description

The system checks data synchronization between the active and standby FlinkServer nodes every 60 seconds. This alarm is generated when the standby FlinkServer node fails to synchronize files with the active FlinkServer node.

This alarm is cleared when the standby FlinkServer synchronizes files with the active FlinkServer.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45641	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

After an active/standby switchover, some configurations may be lost. Some jobs and connections of the FlinkServer are interrupted, but the FlinkServer can still provide services properly.

Possible Causes

- The link between the active and standby FlinkServer nodes is interrupted.
- The synchronization file does not exist or the file permission is required.

Handling Procedure

Check whether the network between the active and standby FlinkServer is in normal state.

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Instance**. View and record the IP addresses of active and standby FlinkServer.

Step 2 Log in to the active FlinkServer node as the **root** user.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Run the following command to check whether the standby FlinkServer is reachable:

ping *IP address of the standby FlinkServer*

- If yes, go to **Step 6**.
- If no, go to **Step 4**.

Step 4 Contact the network administrator to check whether the network is faulty.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check whether the storage space of the /srv/BigData/LocalBackup directory is insufficient.

Step 6 Run the following command to check whether the storage space of the **/srv/BigData/LocalBackup** directory is insufficient:

df -hl /srv/BigData/LocalBackup

- If yes, go to **Step 7**.
- If no, go to **Step 10**.

Step 7 Run the following command to clear unnecessary backup files:

rm -rf *Directory to be cleared*

The following are two examples:

rm -rf /srv/BigData/LocalBackup/0/default-oms_20191211143443

Step 8 On FusionInsight Manager, choose **O&M** > **Backup and Restoration** > **Backup Management**.

In the **Operation** column of the backup task, click **Configure** and change the value of **Maximum Number of Backup Copies** to reduce the number of backup file sets.

Step 9 Wait for 1 minute and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 10**.

Check whether the synchronization file exists and whether the file permission is valid.

Step 10 Run the following command to check whether the synchronization file exists:

find /srv/BigData/ -name "sed*"

find /opt -name "sed*"

- If yes, go to **Step 11**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 12**.

Step 11 Run the following command to check the synchronization file information and permission queried in **Step 10**:

ll *Path of the file you want to search for*

- If the file size is 0 and all values in the permission column are -, the file is a junk file. Run the following command to delete it:

  rm -rf *Files to be deleted*

  Wait for several minutes and check whether the alarm is cleared. If the alarm persists, go to **Step 12**.
- If the file size is not 0, go to **Step 12**.

Step 12 View the log file generated when the alarm is reported.

1. Run the following command to go to the HA run log file path of the current cluster:

   cd /var/log/Bigdata/flink/flinkserver/ha/runlog
2. Decompress log file and view the logs generated when the alarm is reported.

   For example, if the name of the file is **ha.log.2021-03-22_12-00-07.gz**, run the following command:

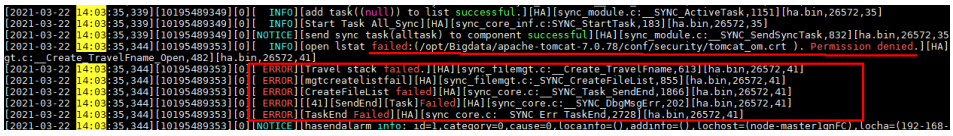
   gunzip *ha.log.2021-03-22_12-00-07.gz*

   vi *ha.log.2021-03-22_12-00-07*

   Check whether error information is displayed before and after the alarm generation time in the logs.

   – If it is displayed, rectify the fault based on the error information. Go to **Step 13**.

   For example, if the following error information is displayed, the directory permission is required. In this case, obtain the directory permission that is the same as the permission on a normal node.

   – If no, go to **Step 14**.

Step 13 Wait for about 10 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 14**.

Collect fault information.

Step 14 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 15 Select FlinkServer information from **Services** and click **OK**.

Step 16 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the hosts to which the role belongs, and click **OK**.

Step 17 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 18 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.392 ALM-45642 RocksDB Continuously Triggers Write Traffic Limiting

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when RocksDB for a job continuously triggers write traffic limiting, that is, the RocksDB write rate is not 0. This alarm is cleared when the RocksDB write rate of the job becomes 0.

The **rocksdb.actual-delayed-write-rate** parameter specifies the RocksDB write rate of a job. Value **0** indicates that the rate is not limited, and other values indicate traffic limiting.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45642	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
JobName	Specifies the job for which the alarm is generated.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

When the rate at which Flink jobs write data to RocksDB is not 0, write traffic limiting is triggered. The possible causes are as follows:

- There are too many MemTables. As a result, write traffic is limited or write stops, and **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold** is generated.
- The size of SST files at level 0 is too large, and **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** is generated.
- The estimated compaction size exceeds the threshold, and **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** is generated.

Handling Procedure

Check whether write traffic limiting or write stop is caused due to too many MemTables.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by following the instructions provided in section **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold**.

Step 4 After ALM-45643 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether write traffic limiting or write stop is caused due to too many SST files at level 0.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 6 In the alarm list, check whether **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** exists.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Handle the alarm by following the instructions provided in section **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold**.

Step 8 After ALM-45644 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check whether write traffic limiting or write stop is caused because the estimated compaction size exceeds the threshold.

Step 9 In the alarm list, check whether **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Handle the alarm by following the instructions provided in section **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold**.

Step 11 After ALM-45647 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

Step 13 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45642 RocksDB Continuously Triggers Write Traffic Limiting**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 14 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 15 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the native Yarn page.

Figure 9-20 Application ID of a job



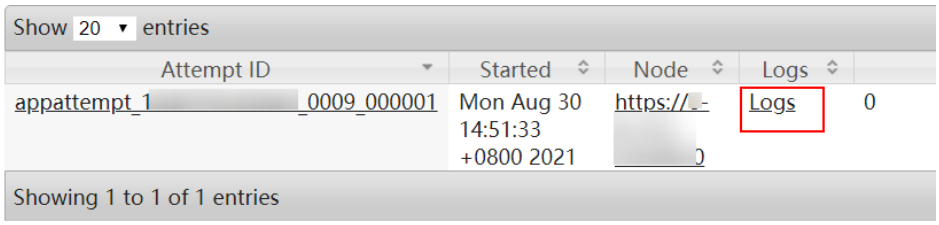
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 16**.
- If no, go to **Step 18**.

Step 16 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

   Figure 9-21 Clicking Logs

2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs.

   Figure 9-22 Clicking the ID in the Attempt ID column

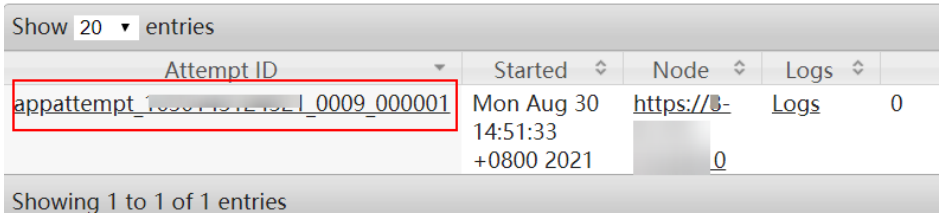
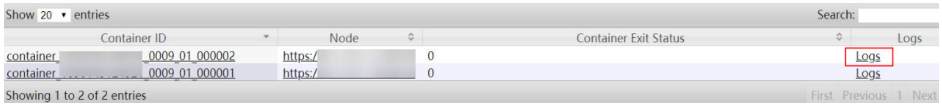
   

   Figure 9-23 Clicking Logs

   NOTE

   You can also log in to Manager as a user who has the FlinkServer management permission. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

Step 17 View the job logs to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 18 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Step 19 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.393 ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the MemTable size of RocksDB for a job continuously exceeds the threshold (**metrics.reporter.alarm.job.alarm.rocksdb.get.micros.threshold**, 50000 microseconds by default). This alarm is cleared when the MemTable size of RocksDB for the job is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45643	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The write pressure of RocksDB is high.

Handling Procedure

Check TaskManager logs for the write pressure of RocksDB and collect logs.

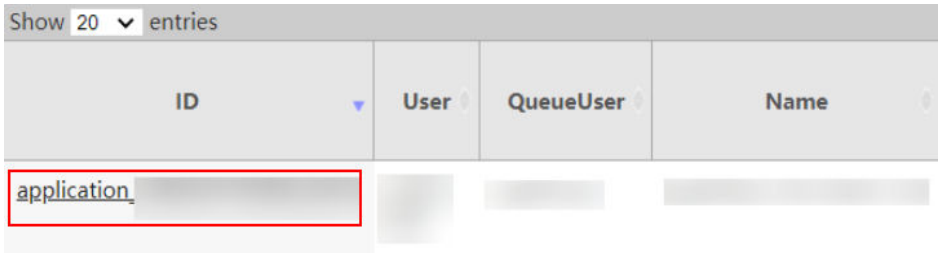
Step 1 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 4 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-24 Application ID of a job

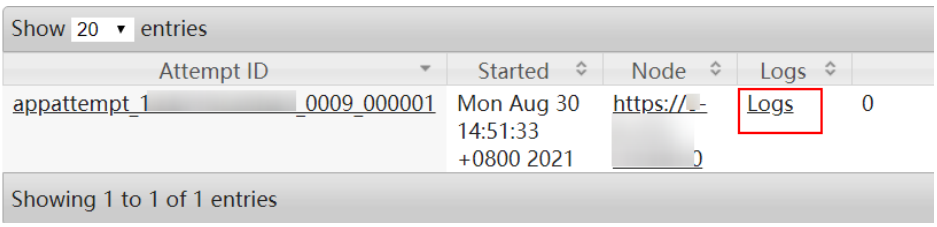


- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

   Figure 9-25 Clicking Logs

2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs. Then go to **Step 7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-26 Clicking the ID in the Attempt ID column

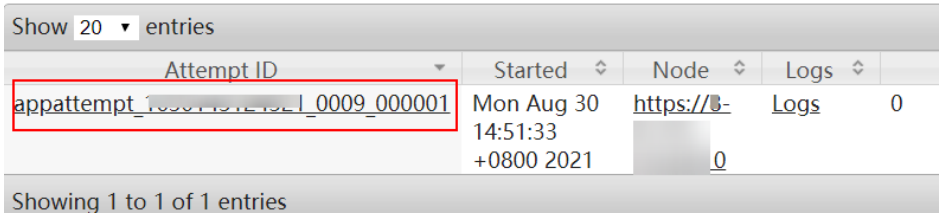
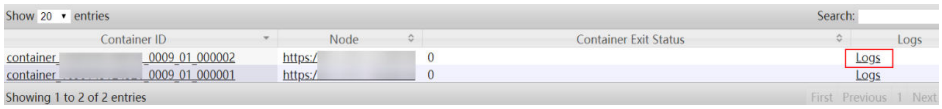


Figure 9-27 Clicking Logs



NOTE

You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Check whether the write pressure of RocksDB is high.

Step 7 Check whether the value of **rocksdb.size-all-mem-tables** (unit: byte) in the TaskManager monitoring logs (keyword **RocksDBMetricPrint**) is greater than or equal to the total write buffer size (Total write buffer = **write_buffer_size** x **max_write_buffer_number**).

- If yes, adjust the values of the following custom parameters on the job development page of the Flink web UI, save the settings, and go to **Step 8**.

Table 9-123 Custom parameters

Parameter	Default Value	Description
state.backend.rocksdb.writebuffer.count	– 2 – **4**: enables **SPINNING_DISK_OPTIMIZED_HIGH_MEM**.	– Number of buffers – **2** to **10** are recommended. Adjust the value based on service requirements.
state.backend.rocksdb.writebuffer.size	64MB	– Buffer size – **64MB** to **256MB** are recommended.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Value	Description
state.backend.rocksdb.thread.num	– 2 – **4**: enables **SPINNING_DISK_OPTIMIZED_HIGH_MEM**.	– Number of flush threads. Increase the number of threads to quickly flush memory data to disks. – When the number of threads is increased, the number of vCores also needs to be increased. – **2** to **10** are recommended.

- If no, go to **Step 9**.

Step 8 Restart the job and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.394 ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the number of SST files at level 0 of RocksDB for a job continuously exceeds the threshold (**state.backend.rocksdb.level0_slowdown_writes_trigger**, 20 by default). This alarm is cleared when the number of SST files at level 0 of RocksDB for the job is less than or equal to the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45644	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

Possible causes are as follows:

- The compaction pressure of RocksDB is too high, and **ALM-45646 Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** and **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** are generated.
- There are too many SST files at level 0.

Handling Procedure

Check whether the compaction pressure of RocksDB is too high and ALM-45646 is generated.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether **ALM-45646 Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Step 3 Handle the alarm by following the instructions provided in section **ALM-45646 Pending Compaction Size of RocksDB Continuously Exceeds the Threshold-Crossing** .

Step 4 After ALM-45646 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the compaction pressure of RocksDB is too high and ALM-45647 is generated.

Step 5 In the alarm list, check whether **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 Handle the alarm by following the instructions provided in section **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold**.

Step 7 After ALM-45647 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether there are too many SST files at level 0.

Step 8 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

Step 9 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 10 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 11 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-28 Application ID of a job



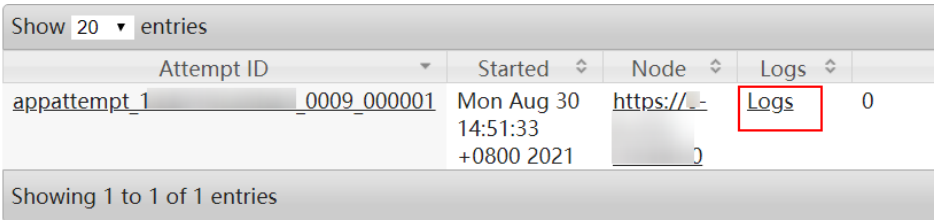
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

   Figure 9-29 Clicking Logs

2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs. Then go to **Step 14**.

   Figure 9-30 Clicking the ID in the Attempt ID column

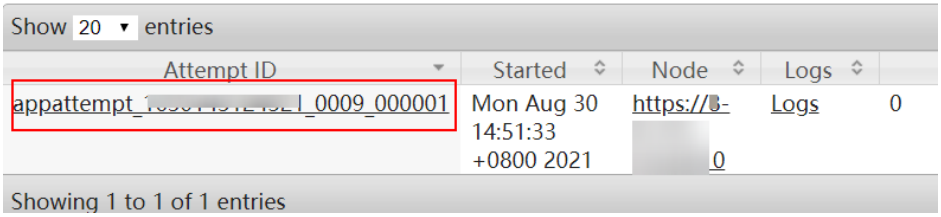
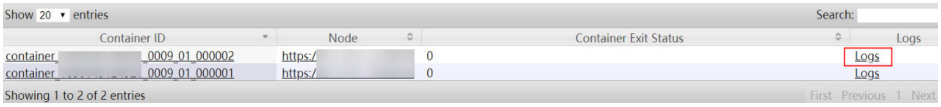
   

   Figure 9-31 Clicking Logs

   NOTE

   You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 13 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Check whether the number of SST files at level 0 is too large.

Step 14 Check whether the value of **rocksdb.num-files-at-level0** in TaskManager monitoring logs (keyword **RocksDBMetricPrint**) is greater than or equal to the value of **state.backend.rocksdb.level0_slowdown_writes_trigger** or **state.backend.rocksdb.level0_stop_writes_trigger**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, modify the values of the following custom parameters on the job development page of the Flink WebUI, save the modification, and go to **Step 15**.

  Table 9-124 Custom parameters

Parameter	Default Value	Description
state.backend.rocksdb.level0_slowdown_writes_trigger	20	– Number of files that trigger slowdown at level 0 – Recommended value: **20** to **30**
state.backend.rocksdb.level0_stop_writes_trigger	36	– Maximum number of files that trigger stop at level 0 – Recommended value: **36** to **46**

- If no, go to **Step 16**.

Step 15 Restart the job and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 16**.

Step 16 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.395 ALM-45645 Pending Flush Size of RocksDB Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the number of pending flush requests of RocksDB for a job continuously reaches *n* times the number of flush/compaction threads. This alarm is cleared when the number of pending flush requests of RocksDB for the job is less than or equal to the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The number of flush/compaction threads is the value of **state.backend.rocksdb.thread.num**. The default value is **2**. If **SPINNING_DISK_OPTIMIZED_HIGH_MEM** is enabled, the default value is **4**.
- The **metrics.reporter.alarm.job.alarm.rocksdb.background.jobs.multiplier** parameter specifies *n* times the number of flush/compaction threads. The default value is **2**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45645	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The number of pending flush requests of RocksDB for the Flink job is too large.

Handling Procedure

Check TaskManager logs for the number of pending flush requests and collect logs.

Step 1 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

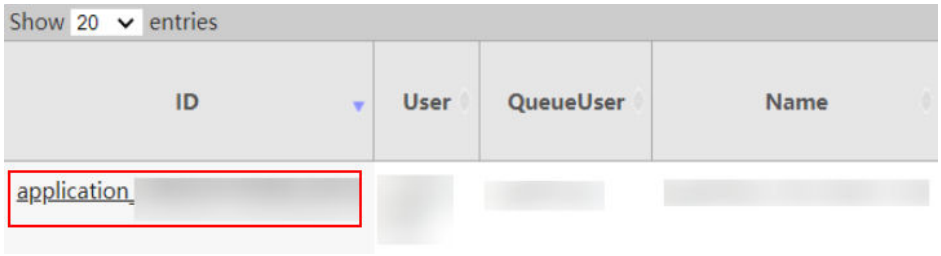
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45645 Pending Flush Size of RocksDB Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 4 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-32 Application ID of a job

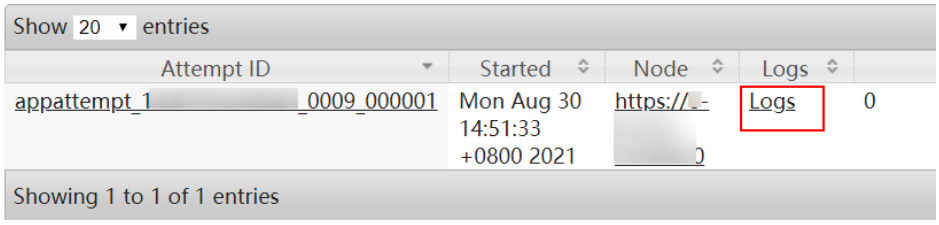


- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

Figure 9-33 Clicking Logs



2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs. Then go to **Step 7**.

Figure 9-34 Clicking the ID in the Attempt ID column

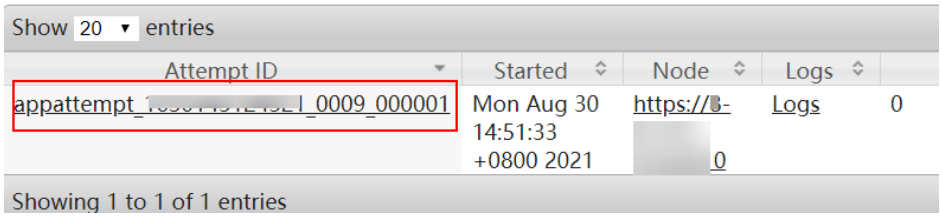
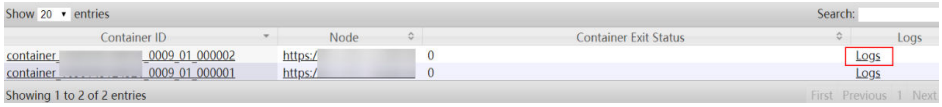


Figure 9-35 Clicking Logs



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Check whether there are too many pending flush requests.

Step 7 Check whether the sum of the values of **rocksdb.mem-table-flush-pending** and **rocksdb.compaction-pending** in TaskManager monitoring logs (keyword **RocksDBMetricPrint**) is greater than *n* times the number of RocksDB threads (**metrics.reporter.alarm.job.alarm.rocksdb.background.jobs.multiplier**, 2 by default). If it is, you can increase the number of RocksDB threads.

- If yes, adjust the values of the following custom parameters on the job development page of the Flink web UI, save the settings, and go to **Step 8**.

  Table 9-125 Custom parameters

Parameter	Default Value	Description
state.backend.rocksdb.thread.num	– 2 – **4**: enables **SPINNING_DISK_OPTIMIZED_HIGH_MEM**.	– Number of flush threads. Increase the number of threads to quickly flush memory data to disks. – When the number of threads is increased, the number of vCores also needs to be increased. – **2** to **10** are recommended.

- If no, go to **Step 9**.

Step 8 Restart the job and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None.

9.11.396 ALM-45646 Pending Compaction Size of RocksDB Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the number of pending compaction requests of RocksDB for a job continuously reaches *n* times the number of flush/compaction threads. This alarm is cleared when the number of pending compaction requests of RocksDB for the job is less than or equal to the threshold.

- The number of flush/compaction threads is the value of **state.backend.rocksdb.thread.num**. The default value is **2**. If **SPINNING_DISK_OPTIMIZED_HIGH_MEM** is enabled, the default value is **4**.
- The **metrics.reporter.alarm.job.alarm.rocksdb.background.jobs.multiplier** parameter specifies *n* times the number of flush/compaction threads. The default value is **2**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45646	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The number of pending compaction requests of RocksDB for the Flink job is too large.

Handling Procedure

Check TaskManager logs for the number of pending compaction requests and collect logs.

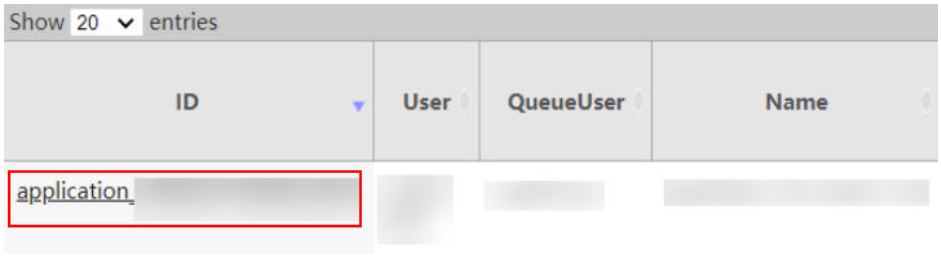
Step 1 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

Step 2 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45646 Pending Compaction Size of RocksDB Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 4 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-36 Application ID of a job

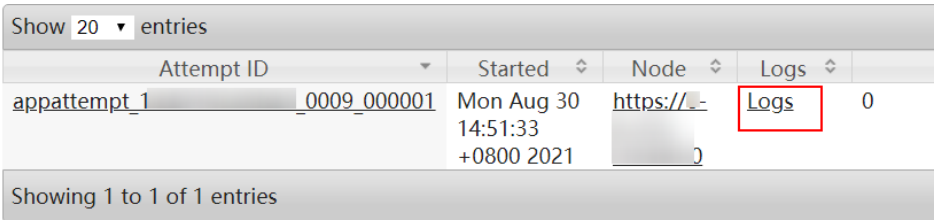


- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

Figure 9-37 Clicking Logs



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs. Then go to **Step 7**.

Figure 9-38 Clicking the ID in the Attempt ID column

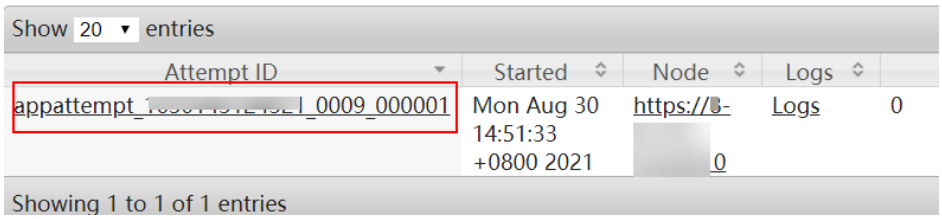
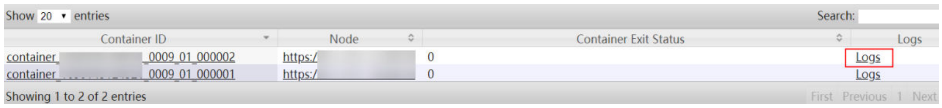


Figure 9-39 Clicking Logs



NOTE

You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Check whether there are too many pending compaction requests.

Step 7 Check whether the sum of the values of **rocksdb.mem-table-flush-pending** and **rocksdb.compaction-pending** in TaskManager monitoring logs (keyword **RocksDBMetricPrint**) is greater than **n** times the number of RocksDB threads (**metrics.reporter.alarm.job.alarm.rocksdb.background.jobs.multiplier**, 2 by default). If it is, you can increase the number of RocksDB threads.

- If yes, adjust the values of the following custom parameters on the job development page of the Flink web UI, save the settings, and go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-126 Custom parameters

Parameter	Default Value	Description
state.backend.rocksdb.thread.num	– 2 – **4**: enables **SPINNING_DISK_OPTIMIZED_HIGH_MEM**.	– Number of flush threads. Increase the number of threads to quickly flush memory data to disks. – When the number of threads is increased, the number of vCores also needs to be increased. – **2** to **10** are recommended.

- If no, go to **Step 9**.

Step 8 Restart the job and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.397 ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the estimated pending compaction size of RocksDB for a job continuously exceeds the threshold. This alarm is cleared when the estimated pending compaction size of RocksDB for the job is less than or equal to the threshold.

The threshold of the estimated pending compaction size is the smaller value of the following two parameters:

- **state.backend.rocksdb.soft-pending-compaction-bytes-limit**. The default value is **64GB**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **state.backend.rocksdb.hard-pending-compaction-bytes-limit**. The default value is **256GB**.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45647	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The estimated compaction data size of RocksDB is too large.

Handling Procedure

Check TaskManager logs for the estimated compaction data size and collect logs.

Step 1 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

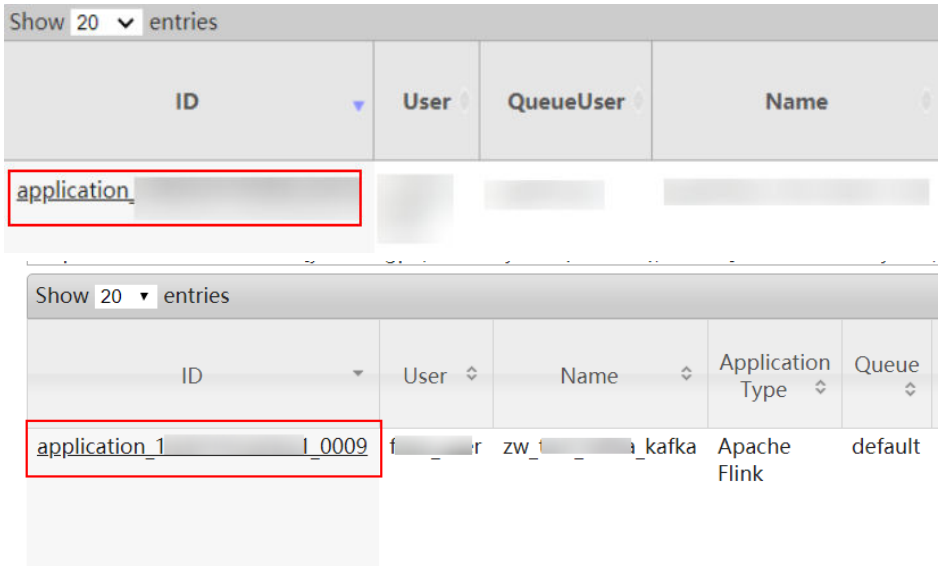
Step 2 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 3 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-40 Application ID of a job

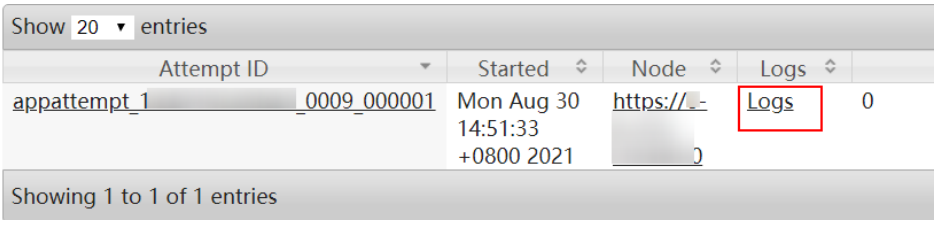


- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Click the application ID of the failed job to go to the job page.

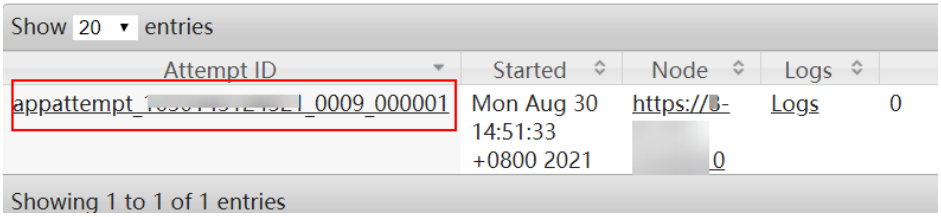
1. Click **Logs** in the **Logs** column to view JobManager logs.

   Figure 9-41 Clicking Logs

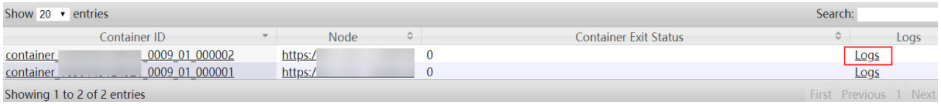
2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs. Then go to **Step 7**.

   Figure 9-42 Clicking the ID in the Attempt ID column

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-43 Clicking Logs



NOTE

You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 6 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Check whether the estimated compaction data size of RocksDB is too large.

Step 7 Check whether the value of **rocksdb.estimate-pending-compaction-bytes** (unit: byte) in TaskManager monitoring logs (keyword **RocksDBMetricPrint**) is greater than or equal to the **soft/hard-pending-compaction** size (values of **state.backend.rocksdb.soft-pending-compaction-bytes-limit** and **state.backend.rocksdb.hard-pending-compaction-bytes-limit**).

- If yes, adjust the values of the following custom parameters on the job development page of the Flink web UI, save the settings, and go to **Step 8**.

Table 9-127 Custom parameters

Parameter	Default Value	Description
state.backend.rocksdb.soft-pending-compaction-bytes-limit	64GB	– When the pending compaction size exceeds the threshold, the write traffic is limited. – **64GB** to **512GB** are recommended.
state.backend.rocksdb.hard-pending-compaction-bytes-limit	256GB	– When the pending compaction size exceeds the threshold, write operations are stopped. – **64GB** to **512GB** are recommended.

- If no, go to **Step 9**.

Step 8 Restart the job and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.398 ALM-45648 RocksDB Frequently Encounters Write-Stopped

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when RocksDB for a job continuously encounters the **is-write-stopped** state. This alarm is cleared when RocksDB for the job no longer or does not continuously encounter the **is-write-stopped** state within an alarm reporting interval.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45648	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The possible causes are as follows:

- There are too many MemTables and **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold** is generated.
- There are too many SST files at level 0, and **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** is generated.
- The estimated compaction size exceeds the threshold, and **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** is generated.

Handling Procedure

Check whether there are too many MemTables.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by following the instructions provided in section **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold**.

Step 4 After ALM-45643 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the number of SST files at level 0 is too large.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 6 In the alarm list, check whether **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Handle the alarm by following the instructions provided in section **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold**.

Step 8 After ALM-45644 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the estimated compaction size exceeds the threshold.

Step 9 In the alarm list, check whether **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Handle the alarm by following the instructions provided in section **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold**.

Step 11 After ALM-45647 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

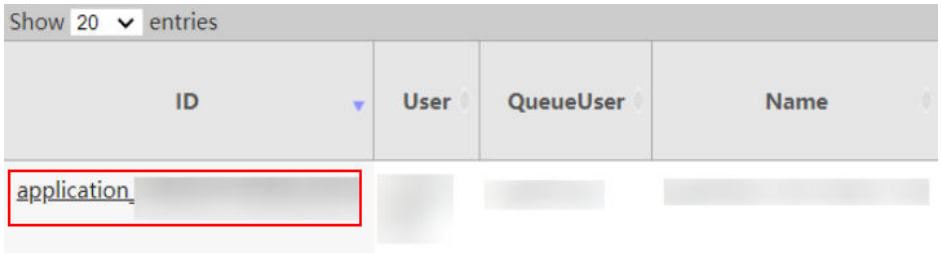
Step 12 Log in to Manager as a user who has the management permission for the current Flink job.

Step 13 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45648 RocksDB Frequently Encounters Write-Stopped**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 14 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 15 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-44 Application ID of a job

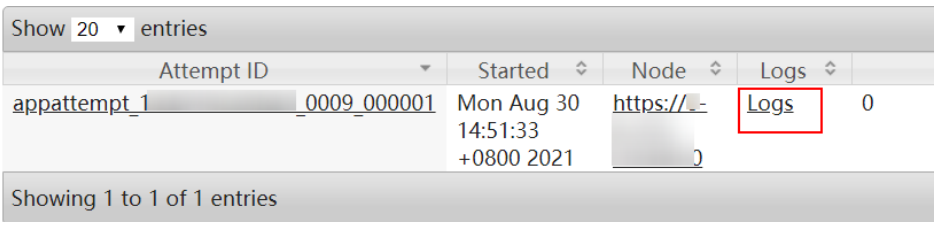


- If yes, go to **Step 16**.
- If no, go to **Step 18**.

Step 16 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

Figure 9-45 Clicking Logs



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs.

Figure 9-46 Clicking the ID in the Attempt ID column

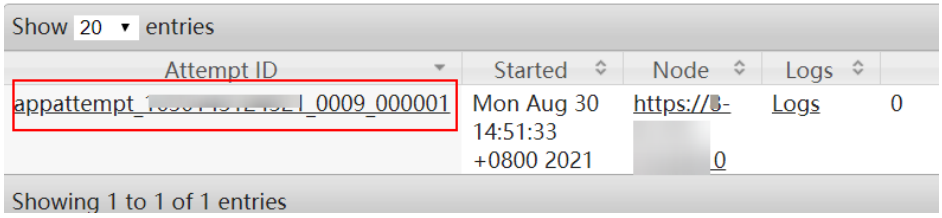
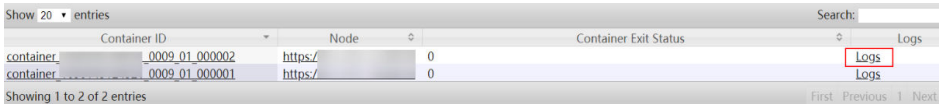


Figure 9-47 Clicking Logs



NOTE

You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

Step 17 View the job logs to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 18 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Step 19 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.399 ALM-45649 P95 Latency of RocksDB Get Requests Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the P95 latency of RocksDB Get requests exceeds the threshold (**metrics.reporter.alarm.job.alarm.rocksdb.get.micros.threshold**, 50000 microseconds by default). This alarm is cleared when the P95 latency of RocksDB Get requests is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45649	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The possible causes are as follows:

- There are too many SST files at level 0, causing slow queries. In addition, **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** is generated.
- The cache hit ratio is lower than 60%, causing frequent swap-ins and swap-outs of the block cache.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check whether the number of SST files at level 0 is too large.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by following the instructions provided in section **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold**.

Step 4 After ALM-45644 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check the cache hit ratio in TaskManager logs and collect logs.

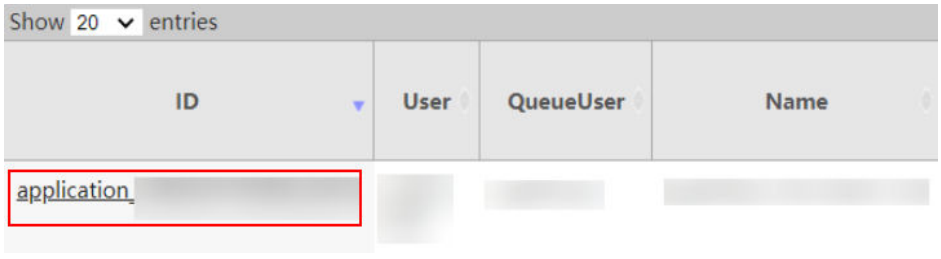
Step 5 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

Step 6 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45649 P95 Latency of RocksDB Get Requests Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 7 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 8 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-48 Application ID of a job



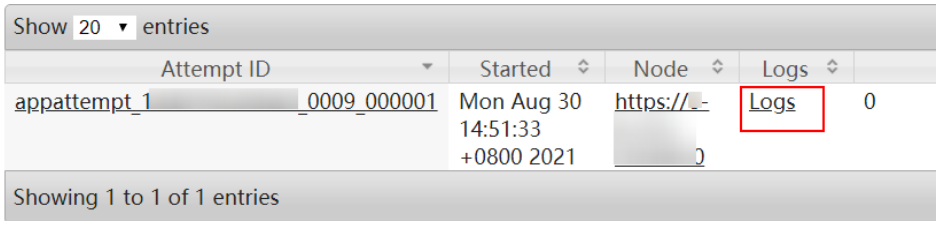
- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 9-49 Clicking Logs



2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs. Then go to **Step 11**.

Figure 9-50 Clicking the ID in the Attempt ID column

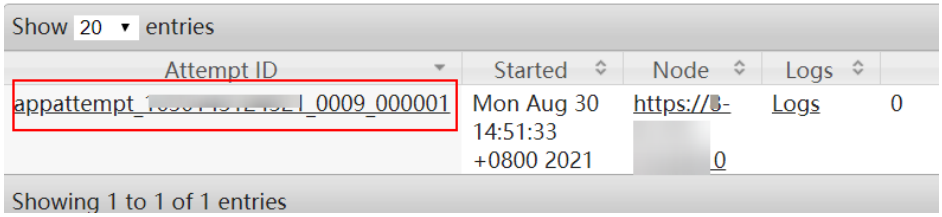
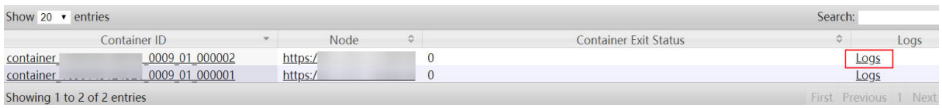


Figure 9-51 Clicking Logs



NOTE

You can also log in to Manager as a user who has the management permission for the current Flink job. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 10 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username*/**bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Check whether the cache hit ratio is too low.

Step 11 Check the values of **rocksdb.block.cache.hit** (cache hit) and **rocksdb.block.cache.miss** (cache miss) in TaskManager monitoring logs (keyword **RocksDBMetricPrint**). Calculate the hit ratio using the following formula and check whether it is less than 60%:

rocksdb.block.cache.hit/(rocksdb.block.cache.hit+rocksdb.block.cache.miss)

- If yes, adjust the values of the following custom parameters on the job development page of the Flink web UI, save the settings, and go to **Step 12**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 9-128 Custom parameters

Parameter	Default Value	Description
state.backend.rocksdb.block.cache-size	– 8MB – **256MB**: enables **SPINNING_DISK_OPTIMIZED_HIGH_MEM**.	– Cache size – **8MB** to **1GB** are recommended.
state.backend.rocksdb.block.blocksize	– 4KB – **128KB**: enables **SPINNING_DISK_OPTIMIZED_HIGH_MEM**.	– Block size – **4KB** to **256KB** are recommended.
state.backend.rocksdb.use-bloom-filter	false	– Whether to speed up indexing. If it is **true**, each new SST file will contain a Bloom filter. – **true** is recommended.

- If no, go to **Step 13**.

Step 12 Restart the job and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Step 13 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.400 ALM-45650 P95 Latency of RocksDB Write Requests Continuously Exceeds the Threshold

This section applies to MRS 3.3.0 or later.

Alarm Description

The system checks the RocksDB monitoring data of jobs at the user-specified alarm reporting interval (**metrics.reporter.alarm.job.alarm.rocksdb.metrics.duration**, 180s by default). This alarm is generated when the P95 latency of RocksDB write requests exceeds

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the threshold (**metrics.reporter.alarm.job.alarm.rocksdb.write.micros.threshold**, 50000 microseconds by default). This alarm is cleared when the P95 latency of RocksDB write requests is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45650	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
ApplicationName	Specifies the name of the application for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
JobName	Specifies the job for which the alarm is generated.

Impact on the System

The checkpoint performance of Flink jobs is affected. There is no impact on the FlinkServer.

Possible Causes

The possible causes are as follows:

- There are too many MemTables. As a result, write traffic is limited or write stops, and **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold** is generated.
- There are too many SST files at level 0, and **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** is generated.
- The estimated compaction size exceeds the threshold, and **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** is generated.

Handling Procedure

Check whether write traffic limiting or write stop is caused due to too many MemTables.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by following the instructions provided in section **ALM-45643 MemTable Size of RocksDB Continuously Exceeds the Threshold**.

Step 4 After ALM-45643 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the number of SST files at level 0 is too large.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 6 In the alarm list, check whether **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Handle the alarm by following the instructions provided in section **ALM-45644 Number of SST Files at Level 0 of RocksDB Continuously Exceeds the Threshold**.

Step 8 After ALM-45644 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check whether the estimated compaction size exceeds the threshold.

Step 9 In the alarm list, check whether **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold** exists.

- If yes, go to **Step 10**.
- If no, go to **Step 12**.

Step 10 Handle the alarm by following the instructions provided in section **ALM-45647 Estimated Pending Compaction Size of RocksDB Continuously Exceeds the Threshold**.

Step 11 After ALM-45647 is cleared, wait a few minutes and check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 12**.

Collect fault information.

Step 12 Log in to FusionInsight Manager as a user who has the FlinkServer management permission.

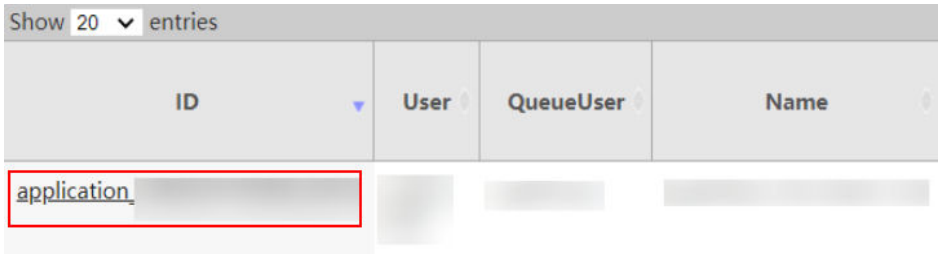
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Choose **O&M** > **Alarm** > **Alarms** > **ALM-45650 P95 Latency of RocksDB Write Requests Continuously Exceeds the Threshold**, view **Location**, and obtain the name of the task for which the alarm is generated.

Step 14 Choose **Cluster** > **Services** > **Yarn** and click the link next to **ResourceManager WebUI** to go to the native Yarn page.

Step 15 Locate the abnormal task based on its name displayed in **Location**, search for and record the application ID of the job, and check whether the job logs are available on the Yarn page.

Figure 9-52 Application ID of a job

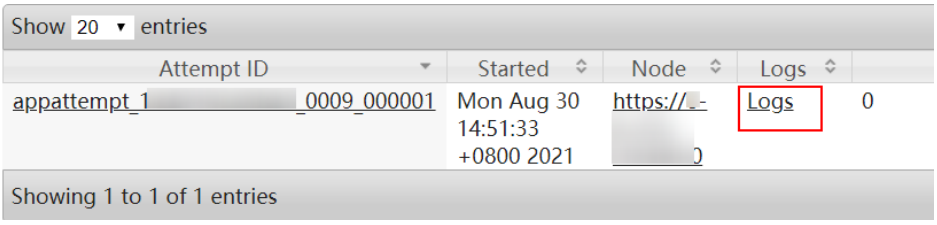


- If yes, go to **Step 16**.
- If no, go to **Step 18**.

Step 16 Click the application ID of the failed job to go to the job page.

1. Click **Logs** in the **Logs** column to view JobManager logs.

Figure 9-53 Clicking Logs



2. Click the ID in the **Attempt ID** column and click **Logs** in the **Logs** column to view and save TaskManager logs.

Figure 9-54 Clicking the ID in the Attempt ID column

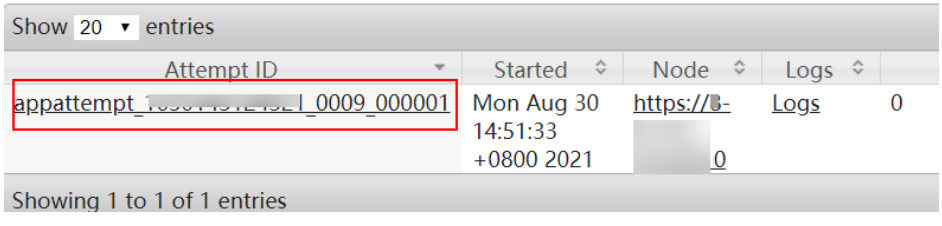
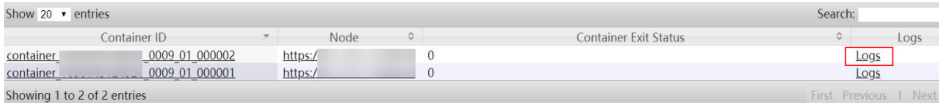


Figure 9-55 Clicking Logs



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

You can also log in to Manager as a user who has the FlinkServer management permission. Choose **Cluster** > **Services** > **Flink**, and click the link next to **Flink WebUI**. On the displayed Flink web UI, click **Job Management**, click **More** in the **Operation** column, and select **Job Monitoring** to view TaskManager logs.

Step 17 View the job logs to rectify the fault, or contact the O&M personnel and send the collected fault logs. No further action is required.

If logs are unavailable on the Yarn page, download logs from HDFS.

Step 18 On Manager, choose **Cluster** > **Services** > **HDFS**, click the link next to **NameNode WebUI** to go to the HDFS page, choose **Utilities** > **Browse the file system**, and download logs in the **/tmp/logs/***Username***/bucket-logs-tfile/***Last four digits of the task application ID/Application ID of the task* directory.

Step 19 View the logs of the failed job to rectify the fault, or contact the O&M personnel and send the collected fault logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.401 ALM-45652 Flink Service Unavailable

This section applies to MRS 3.3.0 or later.

Alarm Description

The alarm module checks the Flink status every 60 seconds. This alarm is generated when the Flink service is unavailable. This alarm is cleared when the Flink service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45652	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the job for which the alarm is generated.

Impact on the System

Flink jobs cannot be submitted with FlinkServer and the Flink client.

Possible Causes

The ZooKeeper, HDFS, Yarn, KrbServer, or DBService service on which Flink depends is unavailable.

Handling Procedure

Check the alarm cause of **Flink Service Unavailable**.

If the alarm cause is ZooKeeper abnormality, check the ZooKeeper service on which Flink depends.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**.

Step 2 In the alarm list, check whether "ALM-13000 ZooKeeper Service Unavailable" exists.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Handle the alarm by referring to **ALM-13000 ZooKeeper Service Unavailable**.

Step 4 After the alarm is cleared, wait a few minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

If the alarm cause is HDFS abnormality, check the HDFS service on which Flink depends.

Step 5 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 6 In the alarm list, check whether "ALM-14000 HDFS Service Unavailable" exists.

- If yes, go to **Step 7**.
- If no, go to **Step 9**.

Step 7 Handle the alarm by referring to **ALM-14000 HDFS Service Unavailable**.

Step 8 After the alarm is cleared, wait a few minutes and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 9**.

If the alarm cause is YARN abnormality, check the YARN service on which Flink depends.

Step 9 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**.

Step 10 In the alarm list, check whether "ALM-18000 Yarn Service Unavailable" exists.

- If yes, go to **Step 11**.
- If no, go to **Step 13**.

Step 11 Handle the alarm by referring to **ALM-18000 Yarn Service Unavailable**.

Step 12 After the alarm is cleared, wait a few minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collect fault information.

Step 13 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 14 Expand the **Service** drop-down list, and select **Flink** for the target cluster.

Step 15 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time respectively. Then, click **Download**.

Step 16 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.402 ALM-45653 Invalid Flink HA Certificate File

This section applies to MRS 3.3.0 or later.

Alarm Description

Flink checks whether the HA certificate file is valid (whether the certificate exists and whether its format is correct) in the first health check or at 01:00:00 every day. This alarm is generated when the certificate file is invalid. This alarm is automatically cleared when the certificate file becomes valid again.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45653	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

FlinkServer in active/standby mode cannot provide services for external systems, and Flink jobs cannot be submitted on the FlinkServer.

Possible Causes

The HA certificate file is invalid.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45653 Invalid Flink HA Certificate File**, view **Location**, obtain the name of the host for which the alarm is generated, and click the host name to view its IP address.

Check whether the HA certificate file in the system is valid.

Step 2 Log in to the host for which the alarm is generated as user **omm**.

Step 3 Run the **cd ${BIGDATA_HOME}/FusionInsight_Flink_*/install/FusionInsight-Flink-*/ha/local/cert** command to go to the directory where the HA certificate is stored.

Step 4 Run the **ls -l** command to check whether the **server.crt** file exists.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Run the **openssl x509 -in server.crt -text -noout** command and check whether the command output is normal.

- If yes, go to **Step 9**.
- If no, go to **Step 6**.

Step 6 Run the **cd ${BIGDATA_HOME}/FusionInsight_Flink_*/install/FusionInsight-Flink-*/flink/sbin** command to go to the Flink script directory.

Step 7 Run the **sh proceed_ha_ssl_cert.sh** command to generate a new HA certificate. Then, check whether the alarm is cleared 1 minute later.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 9**.
- If no, no further action is required.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Flink** for the target cluster.

Step 11 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.403 ALM-45654 Flink HA Certificate Is About to Expire

This section applies to MRS 3.3.0 or later.

Alarm Description

Flink checks whether the HA certificate file is about to expire in the first health check or at 01:00:00 every day. This alarm is generated when the remaining validity period is less than or equal to 30 days. This alarm is automatically cleared when the remaining validity period is greater than 30 days.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45654	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the certificate expires, the HA function of the FlinkServer in active/standby mode is affected. Flink jobs cannot be submitted on the FlinkServer. For FlinkServers in dual-active mode, the HA function is not affected.

Possible Causes

The HA certificate is about to expire.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45654 Flink HA Certificate Is About to Expire**, view **Location**, obtain the name of the host for which the alarm is generated, and click the host name to view its IP address.

Check whether the HA certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the host for which the alarm is generated as user **omm**.

Step 3 Run the **cd ${BIGDATA_HOME}/FusionInsight_Flink_*/install/FusionInsight-Flink-*/ha/local/cert** command to go to the directory where the HA certificate is stored.

Step 4 Run the **openssl x509 -noout -text -in server.crt** command to query the effective time and due time of the HA certificate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Perform **Step 6** to **Step 7** during off-peak hours to update the certificate file as needed.

Step 6 Run the **cd ${BIGDATA_HOME}/FusionInsight_Flink_*/install/FusionInsight-Flink-*/flink/sbin** command to go to the Flink script directory.

Step 7 Run the **sh proceed_ha_ssl_cert.sh** command to generate a new HA certificate. Then, check whether the alarm is cleared 1 minute later.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 8 On the node where the standby FlinkServer instance is located, repeat **Step 6** to **Step 7**. Then, check whether the alarm is cleared 1 minute later.

- If yes, go to **Step 9**.
- If no, go to **Step 10**.

Step 9 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 10**.
- If no, no further action is required.

Collect fault information.

Step 10 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 11 Expand the **Service** drop-down list, and select **Flink** for the target cluster.

Step 12 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 13 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.404 ALM-45655 Flink HA Certificate File Has Expired

This section applies to MRS 3.3.0 or later.

Alarm Description

Flink checks whether the HA certificate file has expired in the first health check or at 01:00:00 every day. This alarm is generated when the HA certificate has expired. This alarm is automatically cleared when the certificate file becomes valid again.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45655	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

FlinkServer in active/standby mode cannot provide services for external systems, and Flink jobs cannot be submitted on the FlinkServer.

Possible Causes

The HA certificate file has expired.

Handling Procedure

View alarm information.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** > **ALM-45655 Flink HA Certificate File Has Expired**, view **Location**, obtain the name of the host for which the alarm is generated, and click the host name to view its IP address.

Check whether the HA certificate file in the system is valid. If it is not, generate a new one.

Step 2 Log in to the host for which the alarm is generated as user **omm**.

Step 3 Run the **cd ${BIGDATA_HOME}/FusionInsight_Flink_*/install/FusionInsight-Flink-*/ha/local/cert** command to go to the directory where the HA certificate is stored.

Step 4 Run the **openssl x509 -noout -text -in server.crt** command to query the effective time and due time of the HA certificate and check whether the HA certificate file is valid.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 9**.
- If no, go to **Step 5**.

Step 5 Run the **cd ${BIGDATA_HOME}/FusionInsight_Flink_*/install/FusionInsight-Flink-*/flink/sbin** command to go to the Flink script directory.

Step 6 Run the **sh proceed_ha_ssl_cert.sh** command to generate a new HA certificate. Then, check whether the alarm is cleared 1 minute later.

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

Step 7 On the node where the standby FlinkServer instance is located, repeat **Step 5** to **Step 6**. Then, check whether the alarm is cleared 1 minute later.

- If yes, go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Check whether this alarm is generated again during periodic system check.

- If yes, go to **Step 9**.
- If no, no further action is required.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Flink** for the target cluster.

Step 11 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.405 ALM-45736 Guardian Service Unavailable

Alarm Description

The alarm module checks the Guardian service status every 60 seconds. This alarm is generated if Guardian is unavailable.

This alarm is cleared after Guardian recovers.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45736	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Guardian cannot work properly, and OBS cannot be accessed.

Possible Causes

- The HDFS service on which the Guardian service depends is abnormal.
- The TokenServer role instance is abnormal.

Handling Procedure

Check the HDFS service status.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. On the page that is displayed, check whether **ALM-14000 HDFS Service Unavailable** is reported.

- If yes, go to **Step 2**.
- If no, go to **Step 3**.

Step 2 Clear this alarm according to the alarm help provided for **ALM-14000 HDFS Service Unavailable**.

After the alarm is cleared, wait a few minutes and check whether the **Guardian Service Unavailable** alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Check all TokenServer instances.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Log in to the node where the TokenServer instance resides as user **omm** and run the **ps -ef|grep "guardian.token.server.Server"** command to check whether the TokenServer process exists on the node.

- If yes, go to **Step 5**.
- If no, restart the faulty TokenServer instance and go to **Step 4**.

Step 4 In the alarm list, check whether the **Guardian Service Unavailable** alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Guardian** for the destination cluster.

Step 7 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.406 ALM-45737 TokenServer Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the TokenServer service every 60 seconds. This alarm is generated when the heap memory usage of the TokenServer instance exceeds the threshold (95% of the maximum memory) for 10 consecutive times.

This alarm is automatically cleared when the system detects that the heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45737	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the heap memory of the Guardian TokenServer instance overflows, OBS cannot be accessed.

Possible Causes

The heap memory of the TokenServer instance is overused or the heap memory is inappropriately allocated.

Handling Procedure

Check heap memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45737 TokenServer Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **TokenServer Heap Memory Usage**. Then click **OK**.

Step 3 Check whether the heap memory used by TokenServer reaches the threshold (95% of the maximum heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, choose **TokenServer** > **Instance Configuration**. Click **All Configurations**, and choose **TokenServer** >

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

System. Set **-Xmx** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for TokenServer cannot meet the heap memory required by the TokenServer process. You are advised to change the value of **-Xmx** in **GC_OPTS** to twice that of the heap memory used by TokenServer. You can change the value based on the actual service scenario. Refer to **Step 2** to view the TokenServer heap memory usage.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During service or instance restart, Guardian may fail to be accessed and jobs cannot access OBS.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Guardian** for the destination cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.407 ALM-45738 TokenServer Direct Memory Usage Exceeds the Threshold

Alarm Description

The system checks the direct memory usage of the TokenServer service every 60 seconds. This alarm is generated when the direct memory usage of the TokenServer instance exceeds the threshold (80% of the maximum memory) for five consecutive times.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is automatically cleared when the system detects that the TokenServer direct memory usage is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45738	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the direct memory of the Guardian TokenServer instance overflows, OBS cannot be accessed.

Possible Causes

The direct memory of the TokenServer process is overused or the direct memory is inappropriately allocated.

Handling Procedure

Check the direct memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45738 TokenServer Direct Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Customize > **CPU and Memory** > **TokenServer Direct Memory Usage**. Then click **OK**.

Step 3 Check whether the direct memory used by TokenServer reaches the threshold (80% of the maximum direct memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, choose **TokenServer** > **Instance Configuration**. Click **All Configurations**, and choose **TokenServer** > **System**. Set **-XX:MaxDirectMemorySize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the direct memory configured for TokenServer cannot meet the direct memory required by the TokenServer process. You are advised to check the direct memory usage of TokenServer and change the value of **-XX:MaxDirectMemorySize** in **GC_OPTS** to the twice of the direct memory used by TokenServer. You can change the value based on the actual service scenario. Refer to **Step 2** to view the TokenServer direct memory usage.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During service or instance restart, Guardian may fail to be accessed and jobs cannot access OBS.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Guardian** for the destination cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.408 ALM-45739 TokenServer Non-Heap Memory Usage Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the TokenServer service every 60 seconds. This alarm is generated when the non-heap memory usage of the TokenServer instance exceeds the threshold (80% of the maximum memory) for five consecutive times.

This alarm is automatically cleared when the system detects that the non-heap memory usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45739	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

If the non-heap memory of the Guardian TokenServer instance overflows, OBS cannot be accessed.

Possible Causes

The non-heap memory of the TokenServer instance is overused or the non-heap memory is inappropriately allocated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check non-heap memory usage.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45739 TokenServer Non-Heap Memory Usage Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, select the role corresponding to the host name of the instance for which the alarm is generated. Click the drop-down list in the upper right corner of the chart area and choose **Customize** > **CPU and Memory** > **TokenServer Non-Heap Memory Usage**. Then click **OK**.

Step 3 Check whether the non-heap memory used by TokenServer reaches the threshold (80% of the maximum non-heap memory by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, choose **TokenServer** > **Instance Configuration**. Click **All Configurations**, and choose **TokenServer** > **System**. Set **-XX: MaxPermSize** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the non-heap memory size configured for the TokenServer instance cannot meet the non-heap memory required by the TokenServer process. You are advised to change the value of **-XX:MaxPermSize** in **GC_OPTS** to twice that of the current non-heap memory size or change the value based on site requirements.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During service or instance restart, Guardian may fail to be accessed and jobs cannot access OBS.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Guardian** for the destination cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.409 ALM-45740 TokenServer GC Duration Exceeds the Threshold

Alarm Description

The system checks the GC duration of the TokenServer process every 60 seconds. This alarm is generated when the GC duration of the TokenServer process exceeds the threshold (12 seconds by default) for five consecutive times.

This alarm is automatically cleared when the system detects that the GC duration is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45740	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

TokenServer responds slowly, and OBS cannot be accessed.

Possible Causes

The heap memory of the TokenServer process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC duration.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms** > **ALM-45740 TokenServer GC Duration Exceeds the Threshold**. Check the location information of the alarm and view the host name of the instance for which the alarm is generated.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, select the role corresponding to the host name of the instance for which the alarm is generated and click the drop-down list in the upper right corner of the chart area. Choose **Customize** > **GC** > **TokenServer GC Duration**. Then click **OK**.

Step 3 Check whether the GC duration of the TokenServer process collected every minute exceeds the threshold (12 seconds by default).

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 On FusionInsight Manager, choose **Cluster** > **Services** > **Guardian**. On the page that is displayed, click the **Instance** tab. On this tab page, choose **TokenServer** > **Instance Configuration**. Click **All Configurations**, and choose **TokenServer** > **System**. Set **-Xmx** in the **GC_OPTS** parameter to a larger value based on site requirements and save the configuration.

NOTE

If this alarm is generated, the heap memory configured for TokenServer cannot meet the heap memory required by the TokenServer process. You are advised to change the value of **-Xmx** in **GC_OPTS** to twice that of the heap memory used by TokenServer. You can change the value based on the actual service scenario. Refer to **Step 2** to view the TokenServer heap memory usage.

Step 5 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

NOTICE

During service or instance restart, Guardian may fail to be accessed and jobs cannot access OBS.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 On FusionInsight Manager, choose **O&M** > **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **Guardian** for the destination cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.410 ALM-45741 Failed to Call the ECS securitykey API

NOTE

This section applies only to MRS 3.3.0 or later.

Alarm Description

Guardian caches the temporary AK/SK of the ECS agency. When the cache does not exist or is about to expire, Guardian calls the securitykey API of ECS to update the AK/SK. This alarm is generated when calling to the API fails.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45741	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The task may fail to obtain the temporary AK/SK for accessing OBS. As a result, OBS cannot be accessed.

Possible Causes

- No ECS agency is bound to the cluster.
- An underlying interface of ECS is abnormal.

Handling Procedure

Check whether an agency is bound to the cluster.

Step 1 Log in to the MRS console.

Step 2 In the navigation pane on the left, choose **Clusters** > **Active Clusters**. On the page that is displayed, click the cluster name to go to its overview page. Then, check whether the cluster is bound to an agency in the O&M management area.

- If yes, go to **4**.
- If no, go to **3**.

Step 3 Click **Select Agency**. On the page that is displayed, rebind the cluster to an agency. Then check whether the alarm is cleared a few minutes later.

- If yes, no further action is required.
- If no, go to **4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Guardian** for the target cluster.

Step 6 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.411 ALM-45742 Failed to Call the ECS Metadata API

Alarm Description

When Guardian calls an IAM API to obtain the temporary AK/SK, it needs to first obtain related metadata via the ECS Metadata API. This alarm is generated when Guardian fails to call the Metadata API.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45742	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The task may fail to obtain the temporary AK/SK for accessing OBS. As a result, OBS cannot be accessed.

Possible Causes

An underlying interface of ECS is abnormal.

Handling Procedure

Collect fault information.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Expand the **Service** drop-down list, and select **Guardian** for the target cluster.

Step 3 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 4 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.412 ALM-45743 Failed to Call the IAM API

Alarm Description

This alarm is generated when Guardian fails to call the IAM API to obtain a temporary AK/SK.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
45743	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

The task may fail to obtain the temporary AK/SK for accessing OBS. As a result, OBS cannot be accessed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The IAM service is abnormal.

Handling Procedure

Collect fault information.

Step 1 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 2 Expand the **Service** drop-down list, and select **Guardian** for the target cluster.

Step 3 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 4 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.413 ALM-47001 MemArtsCC Service Unavailable

Alarm Description

The system checks the status of the ZooKeeper service on which the MemArtsCC depends every 60 seconds. This alarm is generated when the ZooKeeper service is unavailable.

This alarm is automatically cleared when the ZooKeeper service is normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
47001	Critical	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	ServiceName	Specifies the service for which the alarm is generated.
	RoleName	Specifies the role for which the alarm is generated.
	HostName	Specifies the host for which the alarm is generated.

Impact on the System

The service will be unavailable.

Possible Causes

The ZooKeeper service on which the MemArtCC service depends is unavailable.

Handling Procedure

Handling ZooKeeper exceptions

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. On the ZooKeeper dashboard page, check whether the ZooKeeper service is faulty.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Rectify the ZooKeeper fault based on the error information and alarm information and the ZooKeeper help document.

Step 3 After the ZooKeeper service becomes normal, wait for 60 seconds and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **MemArtsCC** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None

9.11.414 ALM-47002 MemArtsCC Disk Fault

Alarm Description

The alarm module checks the status of the local disk used by MemArtsCC every 60 seconds. This alarm is generated when the alarm module detects that the disk status is abnormal. This alarm is cleared when the disk becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
47002	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm is generated.
	ServiceName	Specifies the service for which the alarm is generated.
	RoleName	Specifies the role for which the alarm is generated.
	HostName	Specifies the host for which the alarm is generated.

Impact on the System

MemArtsCC becomes abnormal or its performance deteriorates.

Possible Causes

The disk used by MemArtsCC is damaged or the permission is read-only.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, search for "ALM-47002 MemArtsCC Disk Fault", and locate the abnormal disk directory based on the alarm information.

Step 2 Contact O&M engineers to check whether the disk is faulty.

- If yes, replace the disk, restart the CCSideCar and CCWorker roles of the faulty node, and go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 4**.

Step 3 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **MemArtsCC** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.415 ALM-47003 Memory Usage of the MemArtsCC Worker Process Exceeds the Threshold

Alarm Description

The system checks the memory usage of the CCWorker process of the MemArtsCC component every 30 seconds. This alarm is generated when the memory usage exceeds the threshold.

This alarm is cleared when the system detects that the memory usage of CCWorker process is lower than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
47003	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm is generated.
	ServiceName	Specifies the service for which the alarm is generated.
	RoleName	Specifies the role for which the alarm is generated.
	HostName	Specifies the host for which the alarm is generated.

Impact on the System

The CCWorker process process may restart, which temporarily reduces the cache hit ratio.

Possible Causes

There is more disk space added to CCWorker, there is a sudden increase in service data, or the workload increases significantly when upper-layer computing services (such as Spark, Hive, and HetuEngine) sent many concurrent requests to the MemArtsCC component.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, search for **ALM-47003 Memory Usage of the MemArtsCC Worker Process Exceeds the Threshold**, and locate the abnormal MemArtsCC instance node based on the alarm information. Obtain the alarm threshold in additional information.

Step 2 Choose **Cluster** > **Services** > **MemArtsCC** > **Configurations** > **All Configurations** > **CCWorker (Role)**, search for the **memory_limit** parameter, and check the maximum available memory of the CCWorker instance in the current cluster. Check whether the service concurrency and data volume increase for a long time and the alarm is not automatically cleared.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 This alarm can be ignored temporarily. After the service peak hours, the alarm will be automatically cleared.

Step 4 Increase the maximum available memory in **Step 2** and click Save.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **MemArtsCC** for the target cluster.

Step 8 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Contact O&M personnel and send the collected logs.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.416 ALM-47004 Average Latency of MemArtsCC Worker Read Requests Exceeds the Threshold

Alarm Description

The system checks the average latency of all read requests to the MemArtsCC Worker process component every 30 seconds. This alarm is generated when the average latency exceeds the limit.

This alarm is cleared when the latency of MemArtsCC Worker read requests is lower than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
47004	Major	Yes

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Category	Parameter	Description
	HostName	Specifies the host for which the alarm was generated.

Impact on the System

The cache performance may deteriorate.

Possible Causes

There is a soaring increase of concurrent requests sent by upper-layer computing services (such as Spark, Hive, and HetuEngine). Or, the service load or disk load increases significantly that even a fault occurs.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, search for **ALM-47004 Average Latency of MemArtsCC Worker Read Requests Exceeds the Threshold**, and locate the abnormal MemArtsCC instance node based on the alarm information. Obtain the alarm threshold in additional information.

Step 2 On the **Alarms** page, check whether a disk fault alarm is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Contact O&M engineers to rectify the disk fault. Then, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 4 Ignore the alarm. Check whether the alarm is cleared when service load goes down.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **MemArtsCC** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Contact O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None

9.11.417 ALM-50201 Doris Service Unavailable

Alarm Description

The alarm module checks the Doris service status every 60 seconds. This alarm is generated when the alarm module detects that all FE and BE instances are abnormal.

This alarm is cleared when any FE or BE instance recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50201	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

FusionInsight Manager cannot be used to perform cluster operations on the Doris service, and Doris service functions are unavailable.

Possible Causes

The FE and BE instances are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Restart the Doris service.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Doris**.

Step 2 On the page that is displayed, click **More** and select **Restart Service**. In the displayed dialog box, verify the password and click **OK** to restart the Doris service. After the service is started, go to **Step 3**.

NOTICE

During the restart of the Doris service, the Doris service is unavailable and cannot provide services for external systems. Tasks connected to the Doris service fail.

Step 3 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.418 ALM-50202 FE CPU Usage Exceeds the Threshold

Alarm Description

The system checks the CPU usage of the FE instance every 30 seconds. The CPU usage has a default threshold. This alarm is generated when the CPU usage exceeds the threshold (**95%** by default) for multiple consecutive times (**3** by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when **Trigger Count** is **1** and the CPU usage is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the CPU usage is less than or equal to 85% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50202	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Processes respond slowly or do not work.

Possible Causes

The alarm threshold or alarm trigger count is improperly configured.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, select the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **CPU Usage of FE (FE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

NOTE

Trigger Count specifies how many times the threshold can be hit before an alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.419 ALM-50203 FE Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the FE instance every 30 seconds. This alarm is generated when the memory usage exceeds the threshold (**95%** by default) for multiple consecutive times (**3** by default).

This alarm is cleared when **Trigger Count** is **1** and the memory usage is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the memory usage is less than or equal to 85% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50203	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Processes respond slowly or do not work.

Possible Causes

- The alarm threshold or alarm trigger count is improperly configured.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, select the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **Memory Usage of FE (FE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

NOTE

Trigger Count specifies how many times the threshold can be hit before an alarm is generated.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.420 ALM-50205 BE CPU Usage Exceeds the Threshold

Alarm Description

The system checks the CPU usage of the BE instance every 30 seconds. This alarm is generated when the CPU usage exceeds the threshold (**95%** by default) for multiple consecutive times (**3** by default).

This alarm is cleared when **Trigger Count** is **1** and the CPU usage is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the CPU usage is less than or equal to 85% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50205	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Processes respond slowly or do not work.

Possible Causes

The alarm threshold or alarm trigger count is improperly configured.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **CPU Usage of BE (BE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

NOTE

Trigger Count specifies how many times the threshold can be hit before an alarm is generated.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.421 ALM-50206 BE Memory Usage Exceeds the Threshold

Alarm Description

The system checks the memory usage of the BE instance every 30 seconds. This alarm is generated when the memory usage exceeds the threshold for multiple consecutive times (**3** by default).

This alarm is cleared when **Trigger Count** is **1** and the memory usage is less than or equal to the threshold. This alarm is cleared when **Trigger Count** is greater than 1 and the memory usage is less than or equal to 85% of the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50206	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Processes respond slowly or do not work.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The alarm threshold or alarm trigger count is improperly configured.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **Memory Usage of BE (BE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

NOTE

Trigger Count specifies how many times the threshold can be hit before an alarm is generated.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.422 ALM-50207 Ratio of Connections to the FE MySQL Port to the Maximum Connections Allowed Exceeds the Threshold

Alarm Description

The system checks the number of MySQL port connections every 30 seconds. This alarm is generated when the ratio of the number of current connections to the maximum number of FE port connections exceeds the threshold (95% by default). The maximum number of FE port connections in the current cluster is specified by the **qe_max_connection** parameter. The default value is **1024**.

This alarm is cleared when the number of MySQL port connections on the FE node is less than or equal to the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50207	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Processes respond slowly or do not work.

Possible Causes

- After the MySQL client is connected to Doris, the connection is not closed.
- A large number of services are concurrently connected to Doris.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M**, and click **Alarm** > **Thresholds** in the navigation pane on the left. Click the name of the desired cluster > **Doris** > **Connection** > **FE MySQL Port Connections (FE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

NOTE

If there are a large number of connections, ensure there are only necessary connections. Otherwise, the service performance may be degraded or even the service may be unavailable.

Step 4 Wait for 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.423 ALM-50208 Failures to Clear Historical Metadata Image Files Exceed the Threshold

Alarm Description

The system checks the number of failures to clear historical metadata image files on the FE node every 30 seconds. This alarm is generated when the number of failures exceeds the threshold (**1** by default).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the system detects that the number of failures is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50208	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Doris metadata occupies more and more disk space, which may cause service exceptions.

Possible Causes

The Doris service is abnormal.

Handling Procedure

Check whether the Doris service is normal.

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Doris**.

Step 2 Check whether **Running Status** of the Doris service is **Normal**.

- If yes, go to **Step 4**.
- If no, go to **Step 3**.

Step 3 If the service process is not started, start it first and check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 4**.

Step 4 Check whether other Doris-related alarms are generated in the cluster. If yes, clear them by referring to the alarm help. Then, check whether this alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

You need to manually clear the alarm after the fault is rectified.

Related Information

None.

9.11.424 ALM-50209 Failures to Generate Doris Metadata Image Files Exceed the Threshold

Alarm Description

The system checks the number of failures to generate metadata image files on the FE node every 30 seconds. This alarm is generated when the number of failures exceeds the threshold (**1** by default).

This alarm is cleared when the system detects that the number of failures is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Alarm Type	Service Type	Auto Cleared
50209	Minor	Error handling	Doris	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Category	Parameter	Description
Location Information	Source	Specifies the cluster or system for which the alarm was generated.
	ServiceName	Specifies the service for which the alarm was generated.
	RoleName	Specifies the role for which the alarm was generated.
	HostName	Specifies the host for which the alarm was generated.
Additional Information	Detail	Specifies the alarm triggering condition.

Impact on the System

The non-master FE node cannot receive the latest metadata image file. As a result, the system reliability deteriorates.

Possible Causes

When the checkpoint thread of the Doris FE detects that the FE heap memory usage exceeds 75% (70% for MRS 3.5.0-LTS or later), the thread determines that the write operation to the image file fails.

Handling Procedure

Check the Doris service status.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, view the role name and obtain the IP address of the instance in **Location** of the alarm whose ID is **50209**.

Step 2 Choose **Cluster** > **Services** > **Doris** > **Instances**, click the FE instance for which the alarm is generated, and click the **Chart** tab of the instance.

Select **CPU and Memory** from chart categories on the left, and check whether the value of **FE Heap Memory Usage Rate** exceeds 75%. The value is 70% for MRS 3.5.0-LTS and later versions.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Choose **Cluster** > **Service** > **Doris**, and click the **Configurations** tab. Search for **FE_GC_OPTS** in the search box, and increase the value of **-Xmx**. The default value is **8GB**.

- If this alarm is occasionally generated, increase the value by 50%. If this alarm is frequently generated, double the value.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Click **Save**. Click **Instances**, select an FE instance whose configuration has expired, and choose **More** > **Restart Instance**.

NOTICE

During FE instance restart, the tasks running on FE nodes will fail. The tasks on FE nodes that are not restarted are not affected.

Step 5 After restart, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 8 Click the edit icon in the upper right corner, and select a time span starting 10 minutes before and ending 10 minutes after when the alarm was generated. Then, click **Download** to collect the logs.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

You need to manually clear the alarm after the fault is rectified.

Related Information

None.

9.11.425 ALM-50210 Maximum Compaction Score of All BE Nodes Exceeds the Threshold

Alarm Description

The system checks the maximum compaction score of all BE nodes every 30 seconds. This alarm is generated when the maximum compaction score exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50210	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Query or write may be delayed.

Possible Causes

The number of concurrent service requests is large in the cluster, or the compaction queue is small.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Performance** > **Maximum compaction score of all BE nodes (BE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Step 5 Choose **Cluster** > **Services** > **Doris** > **Configurations** > **All Configurations** > **BE(Role)** > **Customization**, add the **max_base_compaction_threads** parameter to **be.conf** with a value of **10**, and add the **max_cumu_compaction_threads** parameter with a value **20**.

Step 6 Click **Save**. Click **Instances**, select the BE instances whose configuration has expired, click **More**, and select **Restart Instance** to restart the Doris BE instances.

NOTICE

During BE instance restart, the tasks running on BE nodes will fail. The tasks on BE nodes that are not restarted are not affected.

Step 7 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.426 ALM-50211 FE Queue Length of BE Periodic Report Tasks Exceeds the Threshold

Alarm Description

The system checks the queue length of each BE periodic report task on FE every 30 seconds. This alarm is generated when the queue length exceeds the threshold (10 by default). This value indicates the number of report tasks waiting on the master FE node. A large value indicates a poor FE processing capability.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the system detects that the queue length of BE periodic report tasks on FE is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50211	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The processing capability of FE is insufficient, affecting the service query speed.

Possible Causes

The processing capability of the master FE node is insufficient due to a large number of concurrent service requests in the Doris cluster or insufficient memory for FE processes.

Handling Procedure

Check the GC duration.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, view the role name and obtain the IP address of the instance in **Location** of the alarm whose ID is **50211**.

Step 2 Choose **Cluster** > **Services** > **Doris** > **Instances**, click the FE instance for which the alarm is generated, and click the **Chart** tab of the instance.

Select **JVM** from **Chart Category** on the left, and check whether **Accumulated GC duration of the old generation** of the FE process is greater than 3 seconds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > **Services** > **Doris** > **Configurations** > **All Configurations** > **FE(Role)** > **JVM**, and increase the value of **-Xmx** in **FE_GC_OPTS**. The default value is **8GB**.

- If this alarm is generated occasionally, increase the value by 0.5 times. If this alarm is generated frequently, double the parameter value.
- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 5 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Queue** > **Queue Length of BE Periodic Report Tasks on the FE (FE)**.

Step 6 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 7 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 8 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.427 ALM-50212 Accumulated Old-Generation GC Duration of the FE Process Exceeds the Threshold

Alarm Description

The system checks the accumulated old-generation GC duration of the FE process every 30 seconds. This alarm is generated when the accumulated GC duration exceeds the threshold (3000 ms by default).

This alarm is cleared when the system detects that the accumulated old-generation GC duration of the FE process is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50212	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

A long GC duration of the FE process may interrupt the services.

Possible Causes

The heap memory of the FE process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Handling Procedure

Check the GC duration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**. In the alarm list, view the role name and obtain the IP address of the instance in **Location** of the alarm whose ID is **50212**.

Step 2 Choose **Cluster** > **Services** > **Doris** > **Instances**, click the FE instance for which the alarm is generated, and click the **Chart** tab of the instance.

Select **JVM** from **Chart Category** on the left, and check whether **Accumulated GC duration of the old generation** of the FE process is greater than 3 seconds.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Choose **Cluster** > **Services** > **Doris** > **Configurations** > **All Configurations** > **FE(Role)** > **JVM**, and increase the value of **-Xmx** in **FE_GC_OPTS**. The default value is **8G**.

- If this alarm is occasionally generated, increase the value by 0.5 times. If this alarm is frequently generated, double the value.
- In the case of large service volume and high service concurrency, you are advised to add instances.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 30 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.428 ALM-50213 Number of Tasks Queuing in the FE Thread Pool for Interacting with BE Exceeds the Threshold

Alarm Description

The system checks the number of queuing tasks in the FE thread pool for interacting with BE every 30 seconds. This alarm is generated when the number of

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

queuing tasks exceeds the threshold (10 by default). This FE thread pool is the working thread pool of ThriftServer. It is specified by **rpc_port** in the **fe.conf** file and is used to interact with BE.

This alarm is cleared when the system detects that the number of tasks queuing in the FE thread pool for interacting with BE is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50213	Minor	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The read and write of the Doris service slows down.

Possible Causes

There are a large number of concurrent service requests, causing too many queuing tasks.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Queue** > **Number of tasks that are queuing in the thread pool for interaction between the FE and the BE (FE)**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.429 ALM-50214 Number of Tasks Queuing in the FE Thread Pool for Task Processing Exceeds the Threshold

Alarm Description

The system checks the number of queuing tasks in the FE thread pool for processing tasks every 30 seconds. This alarm is generated when the number of queuing tasks exceeds the threshold (10 by default). This thread pool is used by the NIO MySQL Server to process tasks.

This alarm is cleared when the system detects that the number of tasks queuing in the FE thread pool for processing tasks is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50214	Minor	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The task execution of the entire system becomes slow and blocked.

Possible Causes

Large tasks may block the task execution of the queue.

Handling Procedure

Check the execution status of FE tasks.

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **Doris**. Click the **Chart** tab, select **Connection** from **Chart Category** in the left pane, and view the **FE MySQL Port Connections** chart. If the number of connections is large, click **Instances**, select the FE instance, and click the **Chart** tab. Select **CPU and Memory** from **Chart Category** and view the **CPU Usage of FE** chart. If the CPU usage is high, check the **Time** field in FE audit **log /var/log/Bigdata/audit/doris/fe/fe.audit.log** to collect statistics on the average task duration. If the value is also high, the alarm is caused by large concurrent tasks.

Step 2 After connecting to Doris, run the following command to check whether the default value of **queryTimeout** is too large. The default value is **300** seconds.

show variables like 'query_timeout';

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Run the following command to shorten the timeout period based on site requirements to block the tasks that take a long time:

set global query_timeout=*xxx***;**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Queue** > **Queue Length of Query Execution Thread Pool (BE)**.

Step 5 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 6 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 7 Wait 10 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.430 ALM-50215 Longest Duration of RPC Requests Received by Each FE Thrift Method Exceeds the Threshold

Alarm Description

The system checks the longest duration of RPC requests received by each FE Thrift method every 30 seconds. This alarm is generated when the longest duration exceeds the threshold (5000 ms by default).

This alarm is cleared when the longest duration of RPC requests received by each FE Thrift method is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50215	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

A longer RPC duration indicates a higher performance load and slower network request processing, which may cause service congestion.

Possible Causes

- The network has a latency.
- There are too many concurrent large SQL tasks.

Handling Procedure

Step 1 Log in to the host where the faulty node is deployed as user **root** and run **ping** *IP addresses of all Doris nodes* to check whether the peer host can be pinged.

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Contact the network administrator to restore the network.

Step 3 On FusionInsight Manager, choose **Cluster** > **Services** > **Doris**. Click the **Chart** tab, select **Connection** from **Chart Category** in the left pane, and view the **FE MySQL Port Connections** chart. If the number of connections is large, click **Instances**, select the FE instance, and click the **Chart** tab. Select **CPU and Memory** from **Chart Category** and view the **CPU Usage of FE** chart. If the CPU usage is high, check the **Time** field in FE audit **log /var/log/Bigdata/audit/doris/fe/fe.audit.log** to collect statistics on the average task duration. If the value is also high, the alarm is caused by large concurrent tasks.

Step 4 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Performance** > **Longest duration of RPC requests received by each method of the FE thrift interface (FE)**.

Step 5 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 7 Wait 10 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 9 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.431 ALM-50216 Memory Usage of the FE Node Exceeds the Threshold

Alarm Description

The system checks the memory usage of the FE node every 30 seconds. This alarm is generated when the memory usage exceeds the threshold (95% by default).

This alarm is cleared when the memory usage of the FE node falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50216	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Task execution and client connection to the FE are affected.

Possible Causes

The FE heap memory is too small.

Handling Procedure

Check the FE heap memory usage.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, select the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **Memory usage of the FE node (FE)**.

1. Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.
2. Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 2 Log in to the FE node for which the alarm is generated as user **omm**, run the **top** command to check the memory usage of processes, locate the process with high memory usage, and check whether the process belongs to the current service and is running properly.

- If yes, go to **Step 3**.
- If no, isolate or stop the process, or adjust the memory size, and check whether the memory is released.

Step 3 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

- During the restart of the Doris service, the Doris service is unavailable and cannot provide services for external systems. Tasks connected to the Doris service fail.
- During instance restart, the tasks running on the nodes of the instance will fail. The tasks on instance nodes that are not restarted are not affected.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.432 ALM-50217 Heap Memory Usage of the FE Node Exceeds the Threshold

Alarm Description

The system checks the heap memory usage of the FE node every 30 seconds. This alarm is generated when the heap memory usage exceeds the threshold (95% by default).

This alarm is cleared when the heap memory usage of the FE node falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50217	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Task execution and client connection to the FE are affected.

Possible Causes

The FE heap memory is too small.

Handling Procedure

Check heap memory usage.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, select the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **Heap memory usage of the FE node (FE)**.

1. Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.
2. Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **Doris** > **FE** > **Configurations** > **All Configurations**, search for the **FE_GC_OPTS** parameter, increase the value of **-Xmx** as required, click **Save**, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- If this alarm is generated, the heap memory configured for the current Doris instance is not enough for data transmission. You are advised to open the instance monitoring page, display the Doris heap memory resource status monitoring chart, and observe the change trend of the heap memory used by Doris in the monitoring chart. Then change the value of **-Xmx** to twice the current heap memory usage or to another value to meet site requirements.
- When setting the heap memory, you can set **-Xms** and **-Xmx** to approximately the same value to prevent performance deterioration caused by heap size adjustment after each GC.
- The sum of **-Xmx** and **XX:MaxPermSize** cannot be greater than the actual physical memory of the node server.

Step 3 Restart the affected services or instances and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

NOTICE

- During the restart of the Doris service, the Doris service is unavailable and cannot provide services for external systems. Tasks connected to the Doris service fail.
- During instance restart, the tasks running on the nodes of the instance will fail. The tasks on instance nodes that are not restarted are not affected.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.433 ALM-50219 Length of the Queue in the Thread Pool for Query Execution Exceeds the Threshold

Alarm Description

The system checks the length of the waiting queue in the query execution thread pool every 30 seconds. This alarm is generated when the length exceeds the threshold (20 by default).

This alarm is cleared when the length of the waiting queue in the current query execution thread pool is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50219	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The task execution of the entire system becomes slow and blocked.

Possible Causes

Large tasks may block the task execution of the queue.

Handling Procedure

Check the execution status of tasks.

Step 1 On FusionInsight Manager, choose **Cluster** > **Services** > **Doris**. Click the **Chart** tab, select **Connection** from **Chart Category** in the left pane, and view the **FE MySQL**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Port Connections chart. If the number of connections is large, click **Instances**, select the FE instance, and click the **Chart** tab. Select **CPU and Memory** from **Chart Category** and view the **CPU Usage of FE** chart. If the CPU usage is high, check the **Time** field in FE audit **log /var/log/Bigdata/audit/doris/fe/fe.audit.log** to collect statistics on the average task duration. If the value is also high, the alarm is caused by large concurrent tasks.

Step 2 After connecting to Doris, run the following command to check the **queryTimeout** value of the system:

show variables like 'query_timeout';

If the value is too large, run the **set global query_timeout=*xxx*;** command to shorten the timeout interval and block tasks that last for a long time.

Step 3 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Queue** > **Queue Length of Query Execution Thread Pool (BE)**.

Step 4 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 5 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 6 Wait 10 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collect fault information.

Step 7 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 8 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 9 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 10 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.434 ALM-50220 Error Rate of TCP Packet Receiving Exceeds the Threshold

Alarm Description

The system checks the rate of TCP packet receiving errors every 30 seconds. This alarm is generated when the error rate exceeds the threshold (5% by default).

This alarm is cleared when the error rate is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50220	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The task fails or data is lost.

Possible Causes

The network is faulty, so data cannot be sent.

Handling Procedure

Step 1 Log in to the host where the faulty node is deployed as user **root** and run **ping** *IP addresses of all Doris nodes* to check whether the peer host can be pinged.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Contact the network administrator to restore the network.

Step 3 Wait for a while and check whether the alarm is cleared in the alarm list.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.435 ALM-50221 BE Data Disk Usage Exceeds the Threshold

Alarm Description

The system checks the usage of BE data disks every 30 seconds. This alarm is generated when the disk usage exceeds the threshold (95% by default).

This alarm is cleared when the system detects that the disk usage is less than the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50221	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

New data fails to be written, and the task is interrupted.

Possible Causes

- The disk space of the cluster is full.
- Data skew occurs among BE nodes.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, and view the role name and the IP address for the hostname in **Location**.

Step 2 Expand the disk capacity of the node for which the alarm is generated.

Step 3 Go to **Step 4** if the expansion fails or the alarm persists after the expansion.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 6 Expand the **Hosts** drop-down list. In the **Select Host** dialog box that is displayed, select the abnormal host, and click **OK**.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 1 hour ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None.

9.11.436 ALM-50222 Disk Status of a Specified Data Directory on BE Is Abnormal

Alarm Description

The system checks the disk status of a specified data directory on BE every 30 seconds. This alarm is generated when the disk status is not **1** (**1** indicates the normal state and **0** indicates the abnormal state). This alarm is cleared when the disk status of the specified data directory on BE becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50222	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Service data may be unavailable, and data queries on the Doris client may fail.

Possible Causes

- The hard disk is faulty.
- The disk permissions are set incorrectly.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Handling Procedure

Check whether a disk alarm is generated.

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms** and check whether **ALM-12014 Partition Lost** or **ALM-12033 Slow Disk Fault** exists.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Rectify the fault by referring to the handling procedure of **ALM-12014 Partition Lost** or **ALM-12033 Slow Disk Fault**. Then, check whether the alarm is cleared.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Modify disk permissions.

Step 4 Choose **O&M** > **Alarm** > **Alarms** and view **Location** and **Additional Information** of the alarm to obtain the location of the faulty disk.

Step 5 Log in to the node for which the alarm is generated as user **root**. Go to the directory where the faulty disk is located, and run the **ll** command to check whether the permission for the faulty disk is **711** and whether the user is **omm**.

- If yes, go to **Step 7**.
- If no, go to **Step 6**.

Step 6 Modify the permission of the faulty disk. For example, if the faulty disk is **data1**, run the following commands:

chown omm:wheel data1

chmod 711 data1

Step 7 Choose **Cluster** > **Services** > **Doris** > **Instances**, select this BE instance, click **More**, and select **Restart Instance**. Wait 5 minutes and check whether an alarm is generated.

- If no, no further action is required.
- If yes, go to **Step 8**.

NOTICE

During BE instance restart, the tasks running on BE nodes will fail. The tasks on BE nodes that are not restarted are not affected.

Collect fault information.

Step 8 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 9 Expand the **Service** drop-down list, and select **Doris** and **OMS** for the target cluster.

Step 10 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 20 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 11 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.437 ALM-50223 Maximum Memory Required by BE Is Greater Than the Remaining Memory of the Machine

Alarm Description

The system checks whether the maximum memory required by BE is greater than the available memory every 30 seconds. This alarm is generated when the value is not **1** (**1** indicates that the maximum required memory is less than or equal to the available memory, and **0** indicates that the maximum required memory is greater than the available memory).

This alarm is cleared when the maximum required memory is less than or equal to the available memory.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50223	Major	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

A task may fail to apply for memory when running.

Possible Causes

Too much BE node memory has been occupied by other processes, or the maximum memory set for the BE service is too large.

Handling Procedure

Check whether the maximum memory set for the BE node is proper.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **CPU and Memory** > **Relationship between the maximum memory size of the BE and the remaining memory size of the machine (BE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Log in to the BE node for which the alarm is generated as user **omm**, run the **top** command to check the memory usage of processes, locate the process with high memory usage, and check whether the process belongs to the current service and is running properly.

- If yes, go to **Step 6**.
- If no, isolate or stop the process, or adjust the memory size, and check whether the memory is released.

Step 6 Log in to the BE node for which the alarm is generated as user **omm** and run the **free -g** command to check the total memory and remaining memory in the system and estimate the memory usage.

Step 7 On FusionInsight Manager, choose **Cluster** > **Services** > **Doris** > **Configurations** > **All Configurations** > **BE(Role)** > **Memory** and decrease the value of **mem_limit**. This parameter specifies the maximum memory allowed for BE. Then save the modification and restart the BE instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

During BE instance restart, the tasks running on BE nodes will fail. The tasks on BE nodes that are not restarted are not affected.

Step 8 After the BE instance is restarted, wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Collect fault information.

Step 9 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 10 Expand the **Service** drop-down list, and select **Doris** for the target cluster.

Step 11 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 12 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.438 ALM-50224 Failures a Certain Task Type on BE Are Increasing

Alarm Description

The system checks whether the number of failed tasks of a certain type on BE is increasing every 30 seconds. This alarm is generated when the system detects that the value is not **1** (**1** indicates that the number of failed tasks of a certain type does not increase, and **0** indicates that the failed tasks of a certain type are increasing).

This alarm is cleared when the system detects that the number of failed tasks of a certain type on BE does not increase.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50224	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

A task fails to be executed repeatedly in a certain scenario.

Possible Causes

A BE exception may occur. As a result, the number of failed tasks increases in a certain scenario.

Handling Procedure

Check whether the alarm threshold or alarm trigger count is properly configured.

Step 1 Log in to FusionInsight Manager, choose **O&M** > **Alarm** > **Thresholds**, click the name of the desired cluster, and choose **Doris** > **Exception** > **Check whether the number of failed tasks of a certain type increases (BE)**.

Step 2 Click the edit button next to **Trigger Count**, change the number based on site requirements, and click **OK**.

Step 3 Click **Modify** in the **Operation** column, change the alarm threshold based on site requirements, and click **OK**.

Step 4 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 6 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 7 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

You need to manually clear the alarm after the fault is rectified.

Related Information

None.

9.11.439 ALM-50225 FE Instance Fault

Alarm Description

The system checks the FE process status every 30 seconds. This alarm is generated when the value is greater than **0** (**0** indicates that the FE process is normal and **1** indicates that the FE process is abnormal).

This alarm is cleared when the system detects that the FE process becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50225	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The FE instance is unavailable and cannot respond to client requests.

Possible Causes

The FE instance is faulty or restarted.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the alarm list, view the role name and obtain the IP address of the instance in **Location** of the alarm whose ID is **50225**.

Step 2 Choose **Cluster** > **Services** > **Doris** > **Instances**, click the FE instance for which the alarm is generated, and check whether **Running Status** of the instance is **Restoring**.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.440 ALM-50226 BE Instance Fault

Alarm Description

The system checks the BE process status every 30 seconds. This alarm is generated when the value is greater than **0** (**0** indicates that the BE process is normal and **1** indicates that the BE process is abnormal).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the system detects that the BE process becomes normal.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50226	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The BE instance is unavailable and cannot provide the data read and write functions.

Possible Causes

The BE instance is faulty or restarted.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**. In the alarm list, view the role name and obtain the IP address of the instance in **Location** of the alarm whose ID is **50226**.

Step 2 Choose **Cluster** > **Services** > **Doris** > **Instances**, click the BE instance for which the alarm is generated, and check whether **Running Status** of the instance is **Restoring**.

- If yes, go to **Step 3**.
- If no, go to **Step 4**.

Step 3 Wait 2 minutes and check whether the alarm is automatically cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 4**.

Collect fault information.

Step 4 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 5 Expand the **Service** drop-down list, select **Doris** for the target cluster, and click **OK**.

Step 6 Click the edit icon in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 7 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.441 ALM-50401 Number of JobServer Jobs Waiting to Be Executed Exceeds the Threshold

Alarm Description

The system checks the number of jobs submitted to JobServer every 30 seconds. This alarm is generated when the number of jobs to be executed exceeds 800.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50401	Critical (default threshold: 900) Major (default threshold: 800)	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

Too many JobServer jobs are detected in the queue. For details about the queue usage, see the **Additional Information** field of this alarm. The impacts are as follows:

1. When the number of JobServer jobs in the queue reaches the maximum (1000 by default), new jobs cannot be added.
2. Before the number of JobServer jobs in the queue reaches the maximum, new JobServer jobs cannot be submitted quickly. For example, it takes more time (even hours) to submit added jobs or add new jobs to Yarn.

Possible Causes

Too many jobs are submitted instantaneously.

Handling Procedure

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **JobGateway**.

Step 2 Click the **Instances** tab, click **Add Instance**, and add JobServer instances based on the number of submitted jobs.

Step 3 After the instances are added, restart the JobGateway service.

NOTICE

The job functions of JobGateway will become unavailable during the service restart.

Step 4 Wait 5 minutes and check whether the alarm is automatically cleared.

If yes, no further action is required.

If no, go to **Step 5**.

Collect fault information.

Step 5 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 Expand the **Service** drop-down list, and select **JobGateway** for the target cluster.

Step 7 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 8 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.442 ALM-50402 JobGateway Service Unavailable

Alarm Description

The system checks the JobGateway service status every 60 seconds. This alarm is generated when the JobGateway service is abnormal.

This alarm is cleared when the JobGateway service recovers.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50402	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The job functions related to the JobGateway component may be unavailable. For example, jobs cannot be submitted, terminated, or viewed on the MRS management plane, and MRS job management APIs (V2) cannot be called.

Possible Causes

The node where the JobGateway service locates is faulty.

Handling Procedure

Step 1 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **JobGateway**, and click the **Instance** tab. Check for JobServer or JobBalancer instances that are faulty or not started and view the host names of these instances.

Step 2 On the **Alarm** page of FusionInsight Manager, check whether the **NodeAgent Process Is Abnormal** alarm is generated.

- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Check whether the host name in the alarm information is the same as the host name in **Step 1**.

- If yes, go to **Step 4**.
- If no, go to **Step 6**.

Step 4 Clear the alarm by following the instructions provided in **ALM-12006 NodeAgent Process Is Abnormal**.

Step 5 In the alarm list, check whether alarm **JobGateway Service Unavailable** is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **JobGateway** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

None.

9.11.443 ALM-50406 Failure Rate of the JobServer Job Submission API Exceeds the Threshold

Alarm Description

The system checks the failure rate of the JobServer job submission API every 30 seconds. This alarm is generated when the failure rate exceeds the threshold (80% by default).

This alarm is cleared when the failure rate falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50406	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

A job may fail to be submitted, for example, through the REST API.

Possible Causes

The JobServer instance on the node is abnormal.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, expand details of this alarm in the right pane, and obtain **HostName** of the node for which the alarm is generated in **Location**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **JobGateway** > **Instances**.

Step 3 Select the instance for which the alarm is generated, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

Step 4 Check whether the instance running status is normal after the restart.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **JobGateway** for the target cluster.

Step 8 Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.444 ALM-50407 Failure Rate of the JobServer Job Query API Exceeds the Threshold

Alarm Description

The system checks the failure rate of the JobServer job query API every 30 seconds. This alarm is generated when the failure rate exceeds the threshold (80% by default).

This alarm is cleared when the failure rate falls below the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50407	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

A job may fail to be queried, for example, through the REST API.

Possible Causes

The JobServer instance on the node is abnormal.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, expand details of this alarm in the right pane, and obtain **HostName** of the node for which the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **JobGateway** > **Instances**.

Step 3 Select the instance for which the alarm is generated, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

Step 4 Check whether the instance running status is normal after the restart.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **JobGateway** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.445 ALM-50408 Failure Rate of the JobServer Job Termination API Exceeds the Threshold

Alarm Description

The system checks the failure rate of the JobServer job termination API every 30 seconds. This alarm is generated when the failure rate exceeds the threshold (80% by default).

This alarm is cleared when the failure rate falls below the threshold.

Alarm Attributes

Alarm ID	Alarm Severity	Auto Cleared
50408	Critical	Yes

Alarm Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm was generated.
ServiceName	Specifies the service for which the alarm was generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm was generated.
HostName	Specifies the host for which the alarm was generated.

Impact on the System

A job may fail to be terminated, for example, through the REST API.

Possible Causes

The JobServer instance on the node is abnormal.

Handling Procedure

Step 1 On FusionInsight Manager, choose **O&M** > **Alarm** > **Alarms**, expand details of this alarm in the right pane, and obtain **HostName** of the node for which the alarm is generated in **Location**.

Step 2 On FusionInsight Manager, choose **Cluster** > **Services** > **JobGateway** > **Instances**.

Step 3 Select the instance for which the alarm is generated, click **More**, and select **Instance Rolling Restart**.

NOTE

Services may be affected or interrupted during the restart. You are advised to perform the restart during off-peak hours.

Step 4 Check whether the instance running status is normal after the restart.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Choose **O&M** > **Alarm** > **Alarms** and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Collect fault information.

Step 6 On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.

Step 7 Expand the **Service** drop-down list, and select **JobGateway** for the target cluster.

Step 8 Click in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.

Step 9 Send the collected fault logs to O&M personnel for help.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Alarm Clearance

This alarm is automatically cleared after the fault is rectified.

Related Information

None.

9.11.446 ALM-12001 Audit Log Dump Failure (For MRS 2.*x* or Earlier)

Description

Cluster audit logs need to be dumped on a third-party server due to the local historical data backup policy. Audit logs can be successfully dumped if the dump server meets the configuration conditions. This alarm is generated when the audit log dump fails because the disk space of the dump directory on the third-party server is insufficient or a user changes the username, password, or dump directory of the dump server.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12001	Minor	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The system can only store a maximum of 50 dump files locally. If the fault persists on the dump server, the local audit log may be lost.

Possible Causes

- The network connection is abnormal.
- The username, password, or dump directory of the dump server does not meet the configuration conditions.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The disk space of the dump directory is insufficient.

Procedure

Step 1 Check whether the username, password, and dump directory are correct.

1. Check on the dump configuration page of MRS Manager to see if they are correct.
   - If yes, go to **Step 3**.
   - If no, go to **Step 1.2**.
2. Change the username, password, or dump directory, and click **OK**.
3. Wait 2 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Reset the dump rule.

1. On MRS Manager, choose **System** > **Dump Audit Log**.
2. Reset dump rules, set the parameters properly, and click **OK**.
3. Wait 2 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

9.11.447 ALM-12002 HA Resource Abnormal (For MRS 2.*x* or Earlier)

Description

The high availability (HA) software periodically checks the WebService floating IP addresses and databases of Manager. This alarm is generated when the HA software detects that the WebService floating IP addresses or databases are abnormal.

This alarm is cleared when the HA software detects that the floating IP addresses or databases are normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12002	Major	Yes

Parameter

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
RESName	Specifies the resource for which the alarm is generated.

Impact on the System

If the WebService floating IP addresses of Manager are abnormal, users cannot log in to or use Manager. If databases of Manager are abnormal, all core services and related service processes, such as alarms and monitoring functions, are affected.

Possible Causes

- The floating IP address is abnormal.
- The database is abnormal.

Procedure

Step 1 Check the floating IP address status of the active management node.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the host address and resource name of the alarm.
2. Log in to the active management node. Run the following commands to switch the user:

   sudo su - root

   su - omm
3. Go to the **${BIGDATA_HOME}/om-0.0.1/sbin/** directory, run the **status-oms.sh** script to check whether the floating IP address of the active Manager is normal. View the command output, locate the row where **ResName** is **floatip**, and check whether the following information is displayed.

   Example:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
10-10-10-160 floatip Normal Normal Single_active
```

- If yes, go to **Step 2**.
- If no, go to **Step 1.4**.

4. Contact the O&M personnel to check whether the floating IP NIC exists.
   - If yes, go to **Step 2**.
   - If no, go to **Step 1.5**.
5. Contact O&M personnel to rectify the NIC fault.

   Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Check the database status of the active and standby management nodes.

1. Log in to the active and standby management nodes, run the **sudo su - root** and **su - ommdba** commands to switch to user **ommdba**, and run the **gs_ctl query** command to check whether the following information is displayed in the command output.

   Command output of the active management node:

   ```
   Ha state:
   LOCAL_ROLE: Primary
   STATIC_CONNECTIONS: 1
   DB_STATE: Normal
   DETAIL_INFORMATION: user/password invalid
    Senders info:
   No information
    Receiver info:
   No information
   ```

   Command output of the standby management node:

   ```
   Ha state:
   LOCAL_ROLE: Standby
   STATIC_CONNECTIONS: 1
   DB_STATE : Normal
   DETAIL_INFORMATION: user/password invalid
    Senders info:
   No information
    Receiver info:
   No information
   ```

   - If yes, go to **Step 2.3**.
   - If no, go to **Step 2.2**.
2. Contact the O&M personnel to check whether a network fault occurs and rectify the fault.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.448 ALM-12004 OLdap Resource Abnormal (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the Ldap resource in Manager is abnormal.

This alarm is cleared when the Ldap resource in Manager recovers and the alarm handling is complete.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12004	Major	Yes

Parameter

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The Manager authentication services are unavailable and cannot provide security authentication and user management functions for web upper-layer services. Users may be unable to log in to Manager.

Possible Causes

The LdapServer process in Manager is abnormal.

Procedure

Step 1 Check whether the LdapServer process in Manager is normal.

1. Log in to the active management node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Run **ps -ef | grep slapd** to check whether the LdapServer resource process in the **${BIGDATA_HOME}/om-0.0.1/** directory of the configuration file is running properly.

   You can determine that the resource is normal as follows:

   a. Run **sh ${BIGDATA_HOME}/om-0.0.1/sbin/status-oms.sh** and find that **ResHAStatus** of the OLdap process is **Normal**.

   b. Run **ps -ef | grep slapd** and find that the slapd process occupies port 21750.

   – If yes, go to **Step 2**.

   – If no, go to **Step 3**.

Step 2 Run **kill -2** *PID of the LdapServer process* and wait 20 seconds. The HA starts the OLdap process automatically. Check whether the status of the OLdap resource is normal.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.449 ALM-12005 OKerberos Resource Abnormal (For MRS 2.*x* or Earlier)

Description

The alarm module monitors the status of the Kerberos resource in Manager. This alarm is generated when the Kerberos resource is abnormal.

This alarm is cleared when the alarm handling is complete and the Kerberos resource status recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12005	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The authentication services are unavailable and cannot provide security authentication functions for web upper-layer services. Users may be unable to log in to MRS Manager.

Possible Causes

The OLdap resource on which OKerberos depends is abnormal.

Procedure

Step 1 Check whether the OLdap resource on which OKerberos depends is abnormal in Manager.

1. Log in to the active management node.
2. Run the following command to check whether the OLdap resource managed by HA is normal:

   sh ${BIGDATA_HOME}/OMSV100R001C00x8664/workspace0/ha/module/hacom/script/status_ha.sh

   The OLdap resource is normal when the OLdap resource is in the **Active_normal** state on the active node and in the **Standby_normal** state on the standby node.

   - If yes, go to **Step 3**.
   - If no, go to **Step 2**.

Step 2 Resolve the problem by following the instructions in **ALM-12004 OLdap Resource Abnormal (For MRS 2.x or Earlier)**. After the OLdap resource status recovers, check whether the OKerberos resource is normal.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.450 ALM-12006 Node Fault (For MRS 2.*x* or Earlier)

Description

Controller checks the NodeAgent status every 30 seconds. This alarm is generated when Controller fails to receive the status report of a NodeAgent for three consecutive times.

This alarm is cleared when Controller can properly receive the status report of the NodeAgent.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12006	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Services on the node are unavailable.

Possible Causes

The network is disconnected, or the hardware is faulty.

Procedure

Step 1 Check whether the network is disconnected or the hardware is faulty.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the host address of the alarm.
2. Log in to the active management node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Run the following command to check whether the faulty node is reachable:

   ping *IP address of the faulty host*

   a. If yes, go to **Step 2**.

   b. If no, go to **Step 1.4**.

4. Contact the O&M personnel to check whether the network is faulty.
   - If yes, go to **Step 2**.
   - If no, go to **Step 1.6**.
5. Rectify the network fault and check whether the alarm is cleared from the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 1.6**.
6. Contact the O&M personnel to check whether a hardware fault (for example, a CPU or memory fault) occurs on the node.
   - If yes, go to **Step 1.7**.
   - If no, go to **Step 2**.
7. Repair the faulty components and restart the node. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.451 ALM-12007 Process Fault (For MRS 2.*x* or Earlier)

Description

The process health check module checks the process status every 5 seconds. This alarm is generated when the process health check module detects that the process connection status is Bad for three consecutive times.

This alarm is cleared when the process can be connected.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12007	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The service provided by the process is unavailable.

Possible Causes

- The instance process is abnormal.
- The drive space is insufficient.

Procedure

Step 1 Check whether the instance process is abnormal.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the host name and service name of the alarm.
2. On the **Alarms** page, check whether the alarm **ALM-12006 Node Fault (For MRS 2.x or Earlier)** is generated.

   If yes, go to **Step 1.3**.

   If no, go to **Step 1.4**.
3. Handle the alarm by following the instructions in **ALM-12006 Node Fault (For MRS 2.x or Earlier)**.
4. Check whether the installation directory user, user group, and permission of the alarm role are correct. The correct user, user group, and the permission are **omm**, **ficommon**, and **750**, respectively.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 1.5**.
5. Run the following commands to set the permission to **750** and **User:Group** to **omm:ficommon**:

   chmod 750 *<folder_name>*

   chown omm:ficommon *<folder_name>*
6. Wait 5 minutes and check whether the ALM-12007 Process Fault alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Check whether the disk space is insufficient.

1. On the MRS cluster details page, click the alarm management tab and check whether ALM-12017 Insufficient Disk Capacity is generated in the alarm list.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 3**.
2. Handle the alarm by following the instructions in **ALM-12017 Insufficient Disk Capacity (For MRS 2.x or Earlier)**.
3. Wait 5 minutes and check whether the ALM-12017 Insufficient Disk Capacity alarm is cleared.

   If yes, go to **Step 2.4**.

   If no, go to **Step 3**.
4. Wait 5 minutes and check whether the alarm is cleared.

   If yes, no further action is required.

   If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.452 ALM-12010 Manager Heartbeat Interruption Between the Active and Standby Nodes (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the active Manager does not receive any heartbeat signal from the standby Manager within 7 seconds.

This alarm is cleared when the active Manager receives heartbeat signals from the standby Manager.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12010	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Local Manager HA Name	Specifies a local Manager HA.
Peer Manager HA Name	Specifies a peer Manager HA.

Impact on the System

When the active Manager process is abnormal, an active/standby failover cannot be performed, and services are affected.

Possible Causes

The link between the active and standby Manager servers is abnormal.

Procedure

Step 1 Check whether the network between the active and standby Manager servers is normal.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the address of the standby Manager server.
2. Log in to the active management node.
3. Run the following command to check whether the standby Manager is reachable:

   ping *heartbeat IP address of the standby Manager*

   - If yes, go to **Step 2**.
   - If no, go to **Step 1.4**.
4. Contact the O&M personnel to check whether the network is faulty.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Rectify the network fault and check whether the alarm is cleared from the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Log in to all Master nodes in the cluster and run the following commands to find all **sed** files and delete them:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

find /srv/BigData/ -name "sed*"

find /opt -name "sed*"

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

9.11.453 ALM-12011 Data Synchronization Exception Between the Active and Standby Manager Nodes (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the standby Manager fails to synchronize files with the active Manager.

This alarm is cleared when the standby Manager synchronizes files with the active Manager.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12011	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Local Manager HA Name	Specifies a local Manager HA.
Peer Manager HA Name	Specifies a peer Manager HA.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Because the configuration files on the standby Manager are not updated, some configurations will be lost after an active/standby switchover. Manager and some components may not run properly.

Possible Causes

The link between the active and standby Manager nodes is interrupted.

Procedure

Step 1 Check whether the network between the active and standby Manager servers is normal.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the address of the standby Manager server.
2. Log in to the active management node. Run the following command to check whether the standby Manager is reachable:

   ping *IP address of the standby Manager*

   - If yes, go to **Step 2**.
   - If no, go to **Step 1.3**.
3. Contact the O&M personnel to check whether the network is faulty.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2**.
4. Rectify the network fault and check whether the alarm is cleared from the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.454 ALM-12012 NTP Service Abnormal (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the NTP service on the current node fails to synchronize time with the NTP service on the active OMS node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the NTP service on the current node synchronizes time properly with the NTP service on the active OMS node.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12012	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The time on the node is inconsistent with that on other nodes in the cluster. Therefore, some MRS applications on the node may not run properly.

Possible Causes

- The NTP service on the current node cannot start properly.
- The current node fails to synchronize time with the NTP service on the active OMS node.
- The key value authenticated by the NTP service on the current node is inconsistent with that on the active OMS node.
- The time offset between the node and the NTP service on the active OMS node is large.

Procedure

Step 1 Check the NTP service on the current node.

1. Check whether the ntpd process is running on the node using the following method. Log in to the node for which the alarm is generated and run the **sudo su - root** command to switch to user **root**. Then run the following command to check whether the command output contains the ntpd process:

   ps -ef | grep ntpd | grep -v grep

   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.2**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Run **service ntp start** to start the NTP service.
3. Wait 10 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check whether the current node can synchronize time properly with the NTP service on the active OMS node.

1. Check whether the node can synchronize time with the NTP service on the active OMS node based on additional information of the alarm.

   If yes, go to **Step 2.2**.

   If no, go to **Step 3**.
2. Check whether the synchronization with the NTP service on the active OMS node is faulty.

   Log in to the node for which the alarm is generated, run the **sudo su - root** command to switch to user **root**, and run the **ntpq -np** command.

   If an asterisk (*) exists before the IP address of the NTP service on the active OMS node in the command output, the synchronization is in normal state. The command output is as follows:

   ```
   remote refid st t when poll reach delay offset jitter
   ==============================================================================
   *10.10.10.162 .LOCL. 1 u 1 16 377 0.270 -1.562 0.014
   ```

   If there is no asterisk (*) before the IP address of the NTP service on the active OMS node, as shown in the following command output, and the value of **refid** is **.INIT.**, the synchronization is abnormal.

   ```
   remote refid st t when poll reach delay offset jitter
   ==============================================================================
   10.10.10.162 .INIT. 1 u 1 16 377 0.270 -1.562 0.014
   ```

   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
3. Rectify the fault, wait 10 minutes, and then check whether the alarm is cleared.

   An NTP synchronization failure is usually related to the system firewall. If the firewall can be disabled, disable it and then check whether the fault is rectified. If the firewall cannot be disabled, check the firewall configuration policies and ensure that port **UDP 123** is enabled (you need to follow specific firewall configuration policies of each system).

   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Check whether the key value authenticated by the NTP service on the current node is consistent with that on the active OMS node.

Run **cat** to check whether the authentication code whose key value index is 1 is the same as the value of the NTP service on the active OMS node.

- If yes, go to **Step 4.1**.
- If no, go to **Step 5**.

Step 4 Check whether the time offset between the node and the NTP service on the active OMS node is large.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Check whether the time offset is large in additional information of the alarm.
   - If yes, go to **Step 4.2**.
   - If no, go to **Step 5**.
2. On the **Hosts** page, select the host of the node, and choose **More** > **Stop All Roles** to stop all the services on the node.

   If the time on the alarm node is later than that on the NTP service of the active OMS node, adjust the time of the alarm node. After adjusting the time, choose **More** > **Start All Roles** to start the services on the node.

   If the time on the alarm node is earlier than that on the NTP service of the active OMS node, wait until the time offset is due and adjust the time of the alarm node. After adjusting the time, choose **More** > **Start All Roles** to start the services on the node.

   NOTE

   If you do not wait, data loss may occur.
3. Wait 10 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.455 ALM-12014 Device Partition Lost (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the system detects that a partition to which service directories are mounted is lost (because the device is removed or goes offline, or the partition is deleted). The system checks the partition status periodically.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12014	Major	• Yes: MRS 1.9.3.10 and later patch versions • No: MRS 2.*x* and earlier versions

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DirName	Specifies the directory for which the alarm is generated.
PartitionName	Specifies the device partition for which the alarm is generated.

Impact on the System

Service data fails to be written into the partition, and the service system runs abnormally.

Possible Causes

- The disk is removed.
- The disk is offline, or a bad sector exists on the disk.

Procedure

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the row that contains the alarm.

Step 3 In the **Alarm Details** area, obtain the values of **HostName**, **PartitionName**, and **DirName** from **Location**.

Step 4 Check whether the disk corresponding to **PartitionName** on **HostName** is inserted to the correct server slot.

- If yes, go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Contact hardware engineers to remove the faulty disk.

Step 6 Use PuTTY to log in to the **HostName** node where an alarm is reported and check whether there is a line containing **DirName** in the **/etc/fstab** file.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Run the **vi /etc/fstab** command to edit the file and delete the line containing **DirName**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Contact hardware engineers to insert a new disk. For details, see the hardware product document of the relevant model. If the faulty disk is in a RAID group, configure the RAID group. For details, see the configuration methods of the relevant RAID controller card.

Step 9 Wait 20 to 30 minutes (The disk size determines the waiting time), and run the **mount** command to check whether the disk has been mounted to the **DirName** directory.

- If yes, perform **Step 10** for MRS 1.9.3.10 or later. For other versions, clear the alarm. No further action is required.
- If no, perform **Step 11**.

Step 10 Wait 2 minutes and check whether the alarm is automatically cleared.

- If yes, no further action is required.
- If no, perform **Step 11**.

Step 11 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Alarm Clearing

MRS 1.9.3.10 and later patch versions: After the fault is rectified, the system automatically clears the alarm.

MRS 2.*x* and earlier versions: After the fault is rectified, the system does not automatically clear the alarm. You need to clear the alarm.

Reference

None

9.11.456 ALM-12015 Device Partition File System Read-Only (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the system detects that a partition to which service directories are mounted enters the read-only mode (due to a bad sector or a faulty file system). The system checks the partition status periodically.

This alarm is cleared when the system detects that the partition to which service directories are mounted exits from the read-only mode (because the file system is restored to read/write mode, the device is removed, or the device is formatted).

Attribute

Alarm ID	Alarm Severity	Auto Clear
12015	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DirName	Specifies the directory for which the alarm is generated.
PartitionName	Specifies the device partition for which the alarm is generated.

Impact on the System

Service data fails to be written into the partition, and the service system runs abnormally.

Possible Causes

The disk is faulty, for example, a bad sector exists.

Procedure

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the row that contains the alarm.

Step 3 In the **Alarm Details** area, obtain **HostName** and **PartitionName** from **Location**. **HostName** indicates the node for which the alarm is generated, and **PartitionName** indicates the partition of the faulty disk.

Step 4 Contact hardware engineers to check whether the disk is faulty. If the disk is faulty, remove it from the server.

Step 5 After the disk is removed, the system reports ALM-12014 Partition Lost. Handle the alarm by following the instructions in **ALM-12014 Device Partition Lost (For MRS 2.x or Earlier)**. After the handling, the alarm is automatically cleared.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.457 ALM-12016 CPU Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the CPU usage every 30 seconds and compares the check result with the default threshold. The CPU usage has a default threshold. This alarm is generated when the CPU usage exceeds the threshold for several times (configurable, 10 times by default) consecutively.

This alarm is cleared when the average CPU usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12016	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Processes respond slowly or do not work.

Possible Causes

- The alarm threshold or alarm hit number is improperly configured.
- The CPU configuration cannot meet service requirements. The CPU usage reaches the upper limit.

Procedure

Step 1 Check whether the alarm threshold or alarm hit number is properly configured.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to MRS Manager and change the alarm threshold and alarm hit number based on CPU usage.
2. Choose **System** > **Threshold Configuration** > **Device** > **Host** > **CPU** > **CPU Usage** > **CPU Usage** and change the alarm threshold based on the actual CPU usage.
3. Choose **System** > **Threshold Configuration** > **Device** > **Host** > **CPU** > **CPU Usage** > **CPU Usage** and change **hit number** based on the actual CPU usage.

   NOTE

   This option defines the alarm check phase. **Interval** indicates the alarm check period and **hit number** indicates the number of times when the CPU usage exceeds the threshold. An alarm is generated when the CPU usage exceeds the threshold for several times consecutively.
4. Wait 2 minutes and check whether the alarm is automatically cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Expand the system.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the address of the node.
2. Log in to the node for which the alarm is generated.
3. Run **cat /proc/stat | awk 'NR==1'|awk '{for(i=2;i<=NF;i++)j+=$i;print "" 100 - ($5+$6) * 100 / j;}'** to check the system CPU usage.
4. If the CPU usage exceeds the threshold, expand the CPU capacity.
5. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.458 ALM-12017 Insufficient Disk Capacity (For MRS 2.*x* or Earlier)

Description

The system checks the host disk usage every 30 seconds and compares the actual disk usage with the threshold. The disk usage has a default threshold. This alarm is generated if the disk usage exceeds the threshold.

To change the threshold, choose **System** > **Threshold Configuration**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the host disk usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12017	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the disk partition for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Service processes become unavailable.

Possible Causes

The disk configuration cannot meet service requirements. The disk usage reaches the upper limit.

Procedure

Step 1 Log in to MRS Manager and check whether the threshold is appropriate.

1. The default threshold is 90%. You can change the threshold to meet service requirements.
   - If yes, go to **Step 2**.
   - If no, go to **Step 1.2**.
2. Choose **System** > **Threshold Configuration** and change the alarm threshold based on the actual disk usage.
3. Wait 2 minutes and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 2**.

Step 2 Check whether the disk is a system disk.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the host name and disk partition information.
2. Log in to the node for which the alarm is generated.
3. Run the **df -h** command to check the system disk partition usage. Check whether the disk is mounted to any of the following directories by using the disk partition name obtained in **Step 2.1**: **/**, **/boot**, **/home**, **/opt**, **/tmp**, **/var**, **/var/log**, **/boot**, and **/srv/BigData**.
   - If yes, the disk is a system disk. Then go to **Step 3.1**.
   - If no, the disk is not a system disk. Then go to **Step 2.4**.
4. Run the **df -h** command to check the system disk partition usage. Determine the role of the disk based on the disk partition name obtained in **Step 2.1**.
5. Check whether the disk is used by HDFS or Yarn.
   - If yes, expand the disk capacity for the Core node. Then go to **Step 2.6**.
   - If no, go to **Step 4**.
6. Wait 2 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Check whether large files are written to the disk.

1. Run the **find / -xdev -size +500M -exec ls -l {} \;** command to view files larger than 500 MB on the node. Check whether such files are written to the disk.
   - If yes, go to **Step 3.2**.
   - If no, go to **Step 4**.
2. Handle the large files and check whether the alarm is cleared 2 minutes later.
   - If yes, no further action is required.
   - If no, go to **Step 4**.
3. Expand the disk capacity.
4. Wait 2 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.459 ALM-12018 Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the memory usage every 30 seconds and compares the actual memory usage with the threshold. The memory usage has a default threshold. This alarm is generated when the detected memory usage exceeds the threshold.

This alarm is cleared when the host memory usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12018	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Processes respond slowly or do not work.

Possible Causes

Memory configuration cannot meet service requirements. The memory usage reaches the upper threshold.

Procedure

Step 1 Expand the system.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the host address of the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Log in to the node for which the alarm is generated.
3. Run **free -m | grep Mem\: | awk '{printf("%s,", ($3-$6-$7) * 100 / $2)}'** to check the system memory usage.
4. If the memory usage exceeds the threshold, expand the memory capacity.
5. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.460 ALM-12027 Host PID Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the PID usage every 30 seconds and compares the actual PID usage with the default threshold. This alarm is generated when the PID usage exceeds the threshold.

This alarm is cleared when the host PID usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12027	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

No PID is available for new processes and service processes are unavailable.

Possible Causes

Too many processes are running on the node. You need to increase the value of **pid_max**. The system is abnormal.

Procedure

Step 1 Increase the value of **pid_max**.

1. On the MRS cluster details page, click the alarm from the real-time alarm list. In the **Alarm Details** area, obtain the IP address of the host for which the alarm is generated.
2. Log in to the node for which the alarm is generated.
3. Run the **cat /proc/sys/kernel/pid_max** command to check the value of **pid_max**.
4. If the PID usage exceeds the threshold, open the **/etc/sysctl.conf** file and change the value of **kernel.pid_max** to twice the value of **pid_max** queried in **Step 1.3**. If **kernel.pid_max** does not exist, add it to the end of the file.

   For example, change the parameter value to **kernel.pid_max=65536** and run the following command to make the parameter take effect immediately:

   sysctl -p

   NOTE

   The maximum value of **kernel.pid_max** is as follows:
   - 32-bit OS: **32768**
   - 64-bit OS: **4194304** (22nd power of 2)
5. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Check whether the system environment is abnormal.

1. Contact the O&M personnel to check whether the operating system is abnormal.
   - If yes, rectify the operating system fault and go to **Step 2.2**.
   - If no, go to **Step 3**.
2. Wait 5 minutes and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.461 ALM-12028 Number of Processes in the D State on the Host Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system periodically checks the number of D state processes of user **omm** on the host every 30 seconds and compares the number with the threshold. The number of processes in the D state on the host has a default threshold. This alarm is generated when the number of processes in the D state exceeds the threshold.

This alarm is cleared when the number is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12028	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

Excessive system resources are used and the service process responds slowly.

Possible Causes

The host responds slowly to I/O (disk I/O and network I/O) requests and a process is in the D state.

Procedure

Step 1 Check the process that is in the D state.

1. Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the address of the host.
2. Log in to the node for which the alarm is generated.
3. Run the following commands to switch the user:

   sudo su - root

   su - omm
4. Run the following command as user **omm** to view the PID of the process that is in the D state:

   ps -elf | grep -v "\[thread_checkio\]" | awk 'NR!=1 {print $2, $3, $4}' | grep omm | awk -F' ' '{print $1, $3}' | grep D | awk '{print $2}'
5. Check whether the command output is empty.
   - If yes, the service process is running properly. Then go to **Step 1.7**.
   - If no, go to **Step 1.6**.
6. Switch to user **root** and run the **reboot** command to restart the alarm host.

   Restarting the host brings certain risks. Ensure that the service process runs properly after the restart.
7. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.462 ALM-12031 User omm or Password Is About to Expire (For MRS 2.*x* or Earlier)

Description

The system starts at 00:00 every day to check whether user **omm** and the password are about to expire every eight hours. This alarm is generated if the user or password is about to expire in 15 days.

The alarm is cleared when the validity period of user **omm** is changed or the password is reset and the alarm handling is complete.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12031	Minor	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The node trust relationship is unavailable and Manager cannot manage the services.

Possible Causes

User **omm** or the password is about to expire.

Procedure

Step 1 Check whether user **omm** and the password in the system are valid.

1. Log in to the faulty node.
2. Run the following command to view the information about user **omm** and the password:

   chage -l omm

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Check whether the user has expired based on the system message.
   a. View the value of **Password expires** to check whether the password is about to expire.
   b. View the value of **Account expires** to check whether the user is about to expire.

   NOTE

   If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password are about to expire within 15 days.

   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2**.
4. Run the following command to modify the validity period configuration:
   - Run the following command to set a validity period for user **omm**:

     chage -E *'specified date'* **omm**
   - Run the following command to set the number of validity days for user **omm**:

     chage -M *'number of days'* **omm**
5. Check whether the alarm is cleared automatically in the next periodic check.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.463 ALM-12032 User ommdba or Password Is About to Expire (For MRS 2.*x* or Earlier)

Description

The system starts at 00:00 every day to check whether user **ommdba** and the password are about to expire every eight hours. This alarm is generated if the user or password is about to expire in 15 days.

The alarm is cleared when the validity period of user **ommdba** is changed or the password is reset and the alarm handling is complete.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12032	Minor	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The OMS database cannot be managed and data cannot be accessed.

Possible Causes

User **ommdba** or the password is about to expire.

Procedure

Step 1 Check whether user **ommdba** and the password in the system are valid.

1. Log in to the faulty node.
2. Run the following command to view the information about user **ommdba** and the password:

   chage -l ommdba
3. Check whether the user has expired based on the system message.
   a. View the value of **Password expires** to check whether the password is about to expire.
   b. View the value of **Account expires** to check whether the user is about to expire.

   NOTE

   If the parameter value is **never**, the user and password are valid permanently; if the value is a date, check whether the user and password are about to expire within 15 days.

- If yes, go to **Step 1.4**.
- If no, go to **Step 2**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Run the following command to modify the validity period configuration:
   - Run the following command to set a validity period for user **ommdba**:

     chage -E *'specified date'* **ommdba**
   - Run the following command to set the number of validity days for user **ommdba**:

     chage -M *'number of days'* **ommdba**
5. Check whether the alarm is cleared automatically in the next periodic check.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.464 ALM-12033 Slow Disk Fault (For MRS 2.*x* or Earlier)

Description

For MRS 2.*x* or earlier:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - The system runs the **iostat** command every 3 seconds, and detects that the **svctm** value exceeds 1000 ms for 10 consecutive periods within 30 seconds.
  - The system runs the **iostat** command every 3 seconds, and detects that more than 60% of I/O exceeds 150 ms within 300 seconds.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - The system runs the **iostat** command every 3 seconds, and detects that the **svctm** value exceeds 1000 ms for 10 consecutive periods within 30 seconds.
  - The system runs the **iostat** command every 3 seconds, and detects that more than 60% of I/O exceeds 20 ms within 300 seconds.

This alarm is automatically cleared when the preceding conditions have not been met for 15 minutes.

For MRS 1.9.3.10 or later:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 1000 ms within 30 seconds in at least seven collection periods.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 150 ms within 300 seconds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For SSDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 1000 ms within 30 seconds in at least seven collection periods.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 20 ms within 300 seconds.

The collection period is 3 seconds, and the detection period is 30 or 300 seconds. This alarm is automatically cleared when none of the preceding conditions are met for three consecutive detection periods (30 or 300 seconds).

NOTE

For details about how to obtain the related parameters, see **Related Information**.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12033	• Minor: MRS 1.9.3.10 and later patch versions • Major: MRS 2.*x* and earlier versions	Yes

Parameters

Name	Meaning
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Host Name	Specifies the host for which the alarm is generated.
DiskName	Specifies the disk for which the alarm is generated.

Impact on the System

Service performance deteriorates, service processing capabilities become poor, and services may be unavailable.

Possible Causes

The disk is aged or has bad sectors.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the disk status.

Step 1 On the MRS cluster details page, click the alarm from the real-time alarm list. In the **Alarm Details** area, obtain information about the host for which the alarm is generated and information about the faulty disk.

Step 2 Check whether the node for which the alarm is generated is in a virtualization environment.

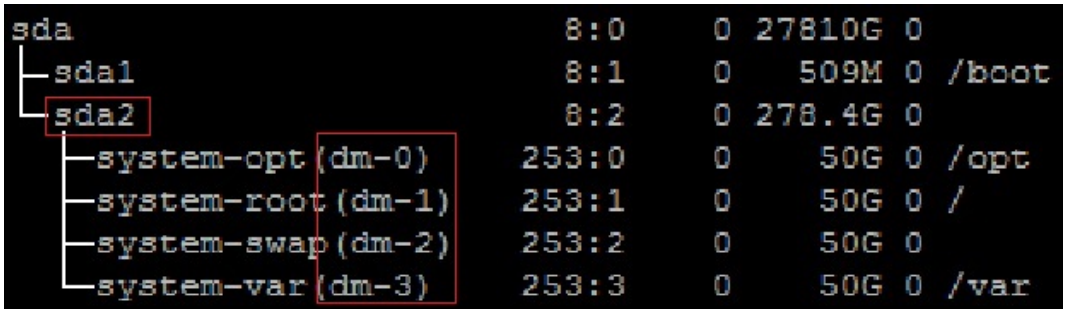
- If yes, go to **Step 3**.
- If no, go to **Step 6**.

Step 3 Check whether the storage performance provided by the virtualization environment meets the hardware requirements. Then, go to **Step 4**.

Step 4 Log in to the alarm node as user **root**, run the **df -h** command, and check whether the command output contains the value of the **DiskName** field.

- If yes, go to **Step 6**.
- If no, go to **Step 5**.

Step 5 Run the **lsblk** command to check whether the mapping between the value of **DiskName** and the disk has been created.



- If yes, go to **Step 6**.
- If no, go to **Step 21**.

Step 6 Log in to the alarm node as user **root**, run the **lsscsi | grep "/dev/sd[x]"** command to view the disk information, and check whether RAID has been set up.

NOTE

In the command, **/dev/sd[x]** indicates the disk name obtained in **Step 1**.

Example:

lsscsi | grep "/dev/sda"

In the command output, if **ATA**, **SATA**, or **SAS** is displayed in the third line, the disk has not been organized into a RAID group. If other information is displayed, RAID has been set up.

- If yes, go to **Step 11**.
- If no, go to **Step 7**.

Step 7 Run the **smartctl -i /dev/sd[x]** command to check whether the hardware supports the SMART tool.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Example:

smartctl -i /dev/sda

In the command output, if "SMART support is: Enabled" is displayed, the hardware supports SMART. If "Device does not support SMART" or other information is displayed, the hardware does not support SMART.

- If yes, go to **Step 8**.
- If no, go to **Step 16**.

Step 8 Run the **smartctl -H --all /dev/sd[x]** command to check basic SMART information and determine whether the disk is working properly.

Example:

smartctl -H --all /dev/sda

Check the value of **SMART overall-health self-assessment test result** in the command output. If the value is **FAILED**, the disk is faulty and needs to be replaced. If the value is **PASSED**, check the value of **Reallocated_Sector_Ct** or **Elements in grown defect list**. If the value is greater than 100, the disk is faulty and needs to be replaced.

- If yes, go to **Step 9**.
- If no, go to **Step 17**.

Step 9 Run the **smartctl -l error -H /dev/sd[x]** command to check the Glist of the disk and determine whether the disk is normal.

Example:

smartctl -l error -H /dev/sda

Check the **Command/Feature_name** column in the command output. If **READ SECTOR(S)** or **WRITE SECTOR(S)** is displayed, the disk has bad sectors. If other errors occur, the disk circuit board is faulty. Both errors indicate that the disk is abnormal and needs to be replaced.

If "No Errors Logged" is displayed, no error log exists. You can perform step 9 to trigger the disk SMART self-check.

- If yes, go to **Step 10**.
- If no, go to **Step 17**.

Step 10 Run the **smartctl -t long /dev/sd[x]** command to trigger the disk SMART self-check. After the command is executed, the time when the self-check is to be completed is displayed. After the self-check is completed, repeat **Step 8** and **Step 9** to check whether the disk is working properly.

Example:

smartctl -t long /dev/sda

- If yes, go to **Step 16**.
- If no, go to **Step 17**.

Step 11 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -H --all /dev/sd[x]** command to check whether the hardware supports SMART.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- In the command, **[sat|scsi]** indicates the disk type. Both types need to be used.
- **[DID]** indicates the slot information. Slots 0 to 15 need to be used.

For example, run the following commands in sequence:

smartctl -d sat+megaraid,0 -H --all /dev/sda

smartctl -d sat+megaraid,1 -H --all /dev/sda

smartctl -d sat+megaraid,2 -H --all /dev/sda

...

Try the command combinations of different disk types and slot information. If "SMART support is: Enabled" is displayed in the command output, the disk supports SMART. Record the parameters of the disk type and slot information when a command is successfully executed. If "SMART support is: Enabled" is not displayed in the command output, the disk does not support SMART.

- If yes, go to **Step 12**.
- If no, go to **Step 15**.

Step 12 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -H --all /dev/sd[x]** command recorded in **Step 11** to check basic SMART information and determine whether the disk is normal.

Example:

smartctl -d sat+megaraid,2 -H --all /dev/sda

Check the value of **SMART overall-health self-assessment test result** in the command output. If the value is **FAILED**, the disk is faulty and needs to be replaced. If the value is **PASSED**, check the value of **Reallocated_Sector_Ct** or **Elements in grown defect list**. If the value is greater than 100, the disk is faulty and needs to be replaced.

- If yes, go to **Step 13**.
- If no, go to **Step 17**.

Step 13 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -l error -H /dev/sd[x]** command to check the Glist of the disk and determine whether the hard disk is working properly.

Example:

smartctl -d sat+megaraid,2 -l error -H /dev/sda

Check the **Command/Feature_name** column in the command output. If **READ SECTOR(S)** or **WRITE SECTOR(S)** is displayed, the disk has bad sectors. If other errors occur, the disk circuit board is faulty. Both errors indicate that the disk is abnormal and needs to be replaced.

If "No Errors Logged" is displayed, no error log exists. You can perform step 9 to trigger the disk SMART self-check.

- If yes, go to **Step 14**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 17**.

Step 14 Run the **smartctl -d [sat|scsi]+megaraid,[DID] -t long /dev/sd[x]** command to trigger the disk SMART self-check. After the command is executed, the time when the self-check is to be completed is displayed. After the self-check is completed, repeat **Step 12** and **Step 13** to check whether the disk is working properly.

Example:

smartctl -d sat+megaraid,2 -t long /dev/sda

- If yes, go to **Step 16**.
- If no, go to **Step 17**.

Step 15 If the configured RAID controller card does not support SMART, the disk does not support SMART. In this case, use the check tool provided by the corresponding RAID controller card vendor to rectify the fault. Then go to **Step 16**.

For example, LSI is a MegaCLI tool.

Step 16 On the alarm details page, click **Clear Alarm**. Check whether the alarm is reported on the same disk again.

If the alarm is reported for more than three times, replace the disk.

- If yes, go to **Step 17**.
- If no, no further action is required.

Replace the disk.

Step 17 On MRS Manager, choose **Alarms**.

Step 18 View the detailed information about the alarm. Check the values of **HostName** and **DiskName** in the location information to obtain the information about the faulty disk for which the alarm is reported.

Step 19 Replace a disk.

Step 20 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 21**.

Collect the fault information.

Step 21 On MRS Manager, choose **System** > **Export Log**.

Step 22 Contact O&M engineers and send the collected logs.

----End

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

Related Information

To obtain the related parameters, perform the following steps:

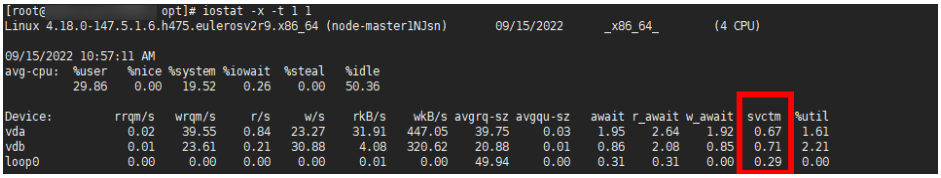
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- For MRS 2.*x* or earlier versions:

  Perform the following operations to detect slow disk faults:

  On the Linux platform, run the **iostat -x -t 1** command to check whether the I/O is faulty. Specifically, check the **svctm** value in the red box in the figure below.

  svctm indicates the I/O service time of the disk.



- For MRS 1.9.3.10 or later patch versions:

  The **svctm** value can be obtained through the following expression:

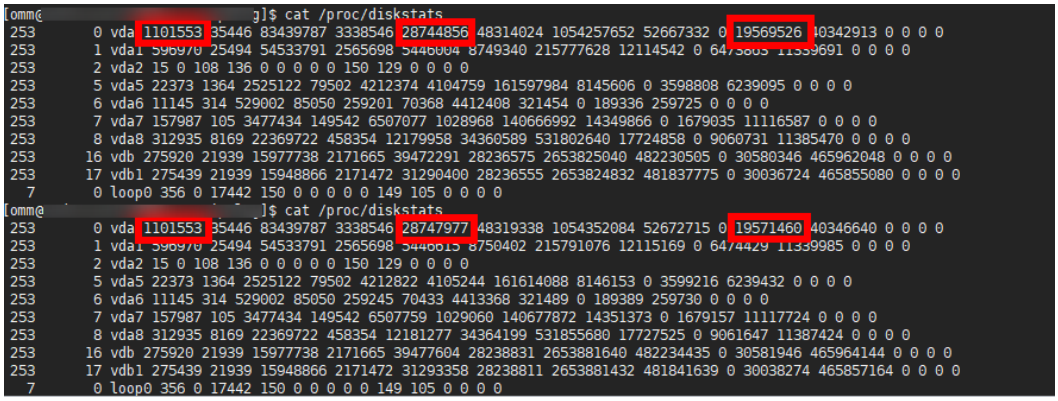
  svctm = (tot_ticks_new - tot_ticks_old) / (rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old)

  When the detection period is 30 seconds, if **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, then **svctm = 0**.

  When the detection period is 300 seconds and **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, if **tot_ticks_new - tot_ticks_old = 0**, then **svctm = 0**; otherwise, the value of **svctm** is infinite.

The parameters in the preceding expression can be obtained as follows:

Obtain the parameter values from the data collected via the **cat /proc/diskstats** command run by the system every 3 seconds. The following shows an example.



In the data collected for the first time, the number in the fourth column is the value of **rd_ios_old**, the number in the eighth column is the value of **wr_ios_old**, and the number in the thirteenth column is the value of **tot_ticks_old**.

In the data collected for the second time, the number in the fourth column is the value of **rd_ios_new**, the number in the eighth column is the value of **wr_ios_new**, and the number in the thirteenth column is the value of **tot_ticks_new**.

In this case, the value of **svctm** is as follows:

(19571460 - 19569526) / (1101553 + 28747977 - 1101553 - 28744856) = 0.6197

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.465 ALM-12034 Periodic Backup Failure (For MRS 2.*x* or Earlier)

Description

This alarm is generated when a periodic backup task fails to be executed. This alarm is cleared when the next backup task is executed successfully.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12034	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
TaskName	Specifies the task name.

Impact on the System

No backup package is available for a long time, so the system cannot be restored in case of exceptions.

Possible Causes

The alarm cause depends on the task details. Handle the alarm according to the logs and alarm details.

Procedure

Checking whether the disk space is insufficient

Step 1 On MRS Manager, choose **Alarms**.

Step 2 In the alarm list, click ⌄ of the alarm and obtain the task name from the **Location** area.

Step 3 Choose **System** > **Back Up Data**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Search for the backup task based on the task name and choose **More** > **View History** in the **Operation** column to view detailed information about the backup task.

Step 5 Choose **Details** > **View** and check whether message "Failed to backup xx due to insufficient disk space, move the data in the /srv/BigData/LocalBackup directory to other directories." exists.

- If yes, go to **Step 6**.
- If no, go to **Step 13**.

Step 6 Choose **Backup Path** > **View** to obtain the backup path.

Step 7 Log in to the node as user **root** and view the mounting details of the node.

df -h

Step 8 Check whether the available space of the node to which the backup path is mounted is less than 20 GB.

- If yes, go to **Step 9**.
- If no, go to **Step 13**.

Step 9 Check whether the backup package exists in the backup directory and whether the available space on the node to which the backup directory is mounted is less than 20 GB.

- If yes, go to **Step 10**.
- If no, go to **Step 13**.

Step 10 Ensure that the available space of the node to which the backup directory is mounted to be greater than 20 GB by moving backup packages out of the backup directory or deleting the backup packages.

Step 11 Start the backup task again and check whether the backup task is executed.

- If yes, go to **Step 12**.
- If no, go to **Step 13**.

Step 12 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 13**.

Collecting fault information

Step 13 On MRS Manager, choose **System** > **Export Log**.

Step 14 Contact O&M engineers and send the collected logs.

----End

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.466 ALM-12035 Unknown Data Status After Recovery Task Failure (For MRS 2.*x* or Earlier)

Description

If a recovery task fails, the system attempts to automatically roll back. If the rollback fails, data may be lost. If this occurs, an alarm is reported. This alarm is cleared when the recovery task is successfully executed later.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12035	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
TaskName	Specifies the task name.

Impact on the System

The data may be lost or the data status may be unknown, which may affect services.

Possible Causes

The possible cause of this alarm is that the component status does not meet the requirements before the restoration task is executed or an error occurs in a step during the restoration task. The error depends on the task details. You can obtain logs and task details to handle the alarm.

Procedure

Checking the component status

Step 1 Log in to MRS Manager and choose **Services**. On the page that is displayed, check whether the running status of the components meets the requirements. (OMS and DBService must be in the normal status, and other components must be stopped.)

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 7**.
- If no, go to **Step 2**.

Step 2 Restore the component status as required and start the recovery task again.

Step 3 Log in to MRS Manager and choose **Alarms**. In the alarm list, click the row containing the alarm and obtain the task name from the **Location** area.

Step 4 Choose **System** > **Recovery Management**. Search for the restoration task based on the task name and view the task details.

Step 5 Start the restoration task and check whether the task is executed.

- If yes, go to **Step 6**.
- If no, go to **Step 7**.

Step 6 After 2 minutes, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Collecting fault information

Step 7 On MRS Manager, choose **System** > **Export Log**.

Step 8 Contact the O&M engineers and send the collected logs.

----End

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

Reference

None

9.11.467 ALM-12037 NTP Server Abnormal (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the NTP server is abnormal.

This alarm is cleared when the NTP server recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12037	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the IP address of the NTP server for which the alarm is generated.

Impact on the System

The NTP server configured on the active OMS node is abnormal. In this case, the active OMS node cannot synchronize time with the NTP server and a time offset may be generated in the cluster.

Possible Causes

- The NTP server network is faulty.
- The NTP server authentication fails.
- The time cannot be obtained from the NTP server.
- The time obtained from the NTP server is not continuously updated.

Procedure

Step 1 Check the NTP server network.

1. On the MRS cluster details page, click the alarm from the real-time alarm list.
2. In the **Alarm Details** area, view the additional information to check whether the NTP server fails to be pinged.
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 2**.
3. Contact the O&M personnel to check the network configuration and ensure that the network between the NTP server and the active OMS node is in normal state. Then, check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Check whether the NTP server authentication fails.

1. Log in to the active management node.
2. Run **ntpq -np** to check whether the NTP server authentication fails. If **refid** of the NTP server is **.AUTH.**, the authentication fails.
   - If yes, go to **Step 5**.
   - If no, go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the time can be obtained from the NTP server.

1. View the alarm additional information to check whether the time cannot be obtained from the NTP server.
   - If yes, go to **Step 3.2**.
   - If no, go to **Step 4**.
2. Contact the O&M personnel to rectify the NTP server fault. After the NTP server is in normal state, check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Check whether the time obtained from the NTP server fails to be updated.

1. View the alarm additional information to check whether the time obtained from the NTP server fails to be updated.
   - If yes, go to **Step 4.2**.
   - If no, go to **Step 5**.
2. Contact the provider of the NTP server to rectify the NTP server fault. After the NTP server is in normal state, check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Reference

None

9.11.468 ALM-12038 Monitoring Indicator Dump Failure (For MRS 2.*x* or Earlier)

Description

This alarm is generated when dumping fails after monitoring indicator dumping is configured on MRS Manager.

This alarm is cleared when dumping is successful.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12038	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The upper-layer management system fails to obtain monitoring indicators from the MRS Manager system.

Possible Causes

- The server cannot be connected.
- The save path on the server cannot be accessed.
- The monitoring indicator file fails to be uploaded.

Procedure

Step 1 Contact the O&M personnel to check whether the network connection between the MRS Manager system and the server is normal.

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Contact the O&M personnel to restore the network and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Choose **System** > **Monitor Dumping Configuration** and check whether the FTP username, password, port, dump mode, and public key configured on the monitoring indicator dumping configuration page are consistent with those on the server.

- If yes, go to **Step 5**.
- If no, go to **Step 4**.

Step 4 Enter the correct configuration, click **OK**, and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Choose **System** > **Monitor Dumping Configuration** and check the configuration items, including the FTP username, save path, and dumping mode.

- If the FTP mode is used, go to **Step 6**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If the SFTP mode is used, go to **Step 7**.

Step 6 Log in to the server. In the default path, check whether the save path (relative path) has the read and write permission on the FTP username.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 7 Log in to the server. In the default path, check whether the save path (absolute path) has the read and write permission on the FTP username.

- If yes, go to **Step 9**.
- If no, go to **Step 8**.

Step 8 Add the read and write permission and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Log in to the server and check whether the save path has sufficient disk space.

- If yes, go to **Step 11**.
- If no, go to **Step 10**.

Step 10 Delete unnecessary files or go to the monitoring indicator dumping configuration page to change the save path. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Step 11 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.469 ALM-12039 GaussDB Data Is Not Synchronized (For MRS 2.*x* or Earlier)

Description

The system checks the data synchronization status between the active and standby GaussDB nodes every 10 seconds. This alarm is generated when the synchronization status cannot be queried for six consecutive times or when the synchronization status is abnormal.

This alarm is cleared when the data synchronization status is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12039	Critical	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.
Local GaussDB HA IP	HA IP address of the local GaussDB.
Peer GaussDB HA IP	HA IP address of the peer GaussDB.
SYNC_PERSENT	Synchronization percentage.

Impact on the System

When data is not synchronized between the active and standby GaussDBs, the data may be lost or abnormal if the active instance becomes abnormal.

Possible Causes

- The network between the active and standby nodes is unstable.
- The standby GaussDB is abnormal.
- The disk space of the standby node is full.

Procedure

Step 1 Go to the MRS cluster details page. In the alarm list on the alarm management tab page, click the row that contains the alarm. In the alarm details, view the IP address of the standby GaussDB node.

Step 2 Log in to the active management node.

Step 3 Run the following command to check whether the standby GaussDB is reachable:

ping *heartbeat IP address of the standby GaussDB*

If yes, go to **Step 6**.

If no, go to **Step 4**.

Step 4 Contact the O&M personnel to check whether the network is faulty.

- If yes, go to **Step 5**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 6**.

Step 5 Rectify the network fault and check whether the alarm is cleared from the alarm list.

- If yes, no further action is required.
- If no, go to **Step 6**.

Step 6 Log in to the standby GaussDB node.

Step 7 Run the following commands to switch the user:

sudo su - root

su - omm

Step 8 Go to the **${BIGDATA_HOME}/om-0.0.1/sbin/** directory.

Run the following command to check whether the resource status of the standby GaussDB is normal:

sh status-oms.sh

In the command output, check whether the following information is displayed in the row where **ResName** is **gaussDB**:

```
10_10_10_231 gaussDB Standby_normal Normal Active_standby
```

- If yes, go to **Step 9**.
- If no, go to **Step 15**.

Step 9 Log in to the standby GaussDB node.

Step 10 Run the following commands to switch the user:

sudo su - root

su - omm

Step 11 Run the **echo ${BIGDATA_DATA_HOME}/dbdata_om** command to obtain the GaussDB data directory.

Step 12 Run the **df -h** command to check the system disk partition usage.

Step 13 Check whether the disk where the GaussDB data directory is mounted is full.

- If yes, go to **Step 14**.
- If no, go to **Step 15**.

Step 14 Contact the O&M personnel to expand the disk capacity. After capacity expansion, wait 2 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Step 15 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.470 ALM-12040 Insufficient System Entropy (For MRS 2.*x* or Earlier)

Description

The system checks the entropy at 00:00:00 every day and performs five consecutive checks each time. First, the system checks whether the rng-tools tool is enabled and correctly configured. If not, the system checks the current entropy. This alarm is generated if the entropy is less than 500 in the five checks.

This alarm is cleared if the true random number mode is configured, random numbers are configured in pseudo-random number mode, or neither the true random number mode nor the pseudo-random number mode is configured but the entropy is greater than or equal to 500 in at least one check among the five checks.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12040	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Decryption failures occur and functions related to decryption are affected, for example, DBService installation.

Possible Causes

The rngd service is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Go to the cluster details page and choose **Alarms**.

Step 2 View the alarm details to obtain the value of the **HostName** field in **Location**.

Step 3 Log in to the node for which the alarm is generated and run the **sudo su - root** command to switch to user **root**.

Step 4 Run the **/bin/rpm -qa | grep -w "rng-tools"** command. If the command is executed successfully, run the **ps -ef | grep -v "grep" | grep rngd | tr -d " " | grep "\-o/dev/random" | grep "\-r/dev/urandom"** command and view the command output.

- If the command is executed successfully, the rngd service is installed, correctly configured, and is running properly. Go to **Step 8**.
- If the command is not executed successfully, the rngd service is not running properly. Then go to **Step 5**.

Step 5 Run the following command to start the rngd service:

echo 'EXTRAOPTIONS="-r /dev/urandom -o /dev/random"' >> /etc/sysconfig/rngd

service rngd start

Step 6 Run the **service rngd status** command to check whether the rngd service is in the running state.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 7 Wait until 00:00:00 when the system checks the entropy again. Check whether the alarm is cleared automatically.

- If yes, no further action is required.
- If no, go to **Step 8**.

Step 8 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.471 ALM-12041 Permission of Key Files Is Abnormal (For MRS 2.*x* or Earlier)

Description

The system checks the permission, users, and user groups of key directories or files every hour. This alarm is generated if any of these is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared after the problem that causes abnormal permission, users, or user groups is solved.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12041	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PathName	Specifies the file path or file name.

Impact on the System

System functions are unavailable.

Possible Causes

The user has manually modified the file permission, user information, or user groups, or the system has experienced an unexpected power-off.

Procedure

Step 1 Check the file permission.

1. Go to the MRS cluster details page and choose **Alarms**.
2. In the details of the alarm, query the **HostName** (name of the alarmed host) and **PathName** (path or name of the involved file).
3. Log in to the alarmed node.
4. Run the **ll** *PathName* command to query the current user, permission, and user group of the file or path.
5. Go to the **${BIGDATA_HOME}/nodeagent/etc/agent/autocheck** directory and run the **vi keyfile** command. Search for the name of the involved file and query the correct permission of the file.
6. Compare the actual permission of the file with the permission obtained in **Step 1.5**. If they are different, change the actual permission, user information, and user group to the correct values.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

7. Wait until the next system check is complete and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.472 ALM-12042 Key File Configurations Are Abnormal (For MRS 2.*x* or Earlier)

Description

The system checks key file configurations every hour. This alarm is generated if any key configuration is abnormal.

This alarm is cleared after the configuration becomes normal.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12042	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PathName	Specifies the file path or file name.

Impact on the System

Functions related to the file are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

The user has manually modified the file configurations or the system has experienced an unexpected power-off.

Procedure

Step 1 Check the file configurations.

1. Go to the MRS cluster details page and choose **Alarms**.
2. In the details of the alarm, query the **HostName** (name of the alarmed host) and **PathName** (path or name of the involved file).
3. Log in to the alarmed node.
4. Manually check and modify the file configurations according to the criteria in **Related Information**.
5. Wait until the next system check is complete and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

- **Checking /etc/fstab**

  Check whether partitions configured in **/etc/fstab** exist in **/proc/mounts** and whether swap partitions configured in **/etc/fstab** match those in **/proc/swaps**.

- **Checking /etc/hosts**

  Run the **cat /etc/hosts** command. If any of the following situations exists, the file configurations are abnormal.
  - The **/etc/hosts** file does not exist.
  - The host name is not configured in the file.
  - The IP address of the host is duplicate.
  - The IP address of the host does not exist in the **ipconfig** list.
  - An IP address in the file is used by multiple hosts.

9.11.473 ALM-12043 DNS Parsing Duration Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the DNS parsing duration every 30 seconds. This alarm is generated when the DNS parsing duration exceeds the threshold (the default threshold is 20,000 ms) for multiple times (the default value is **2**).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Status** > **DNS Resolution Duration** > **DNS Resolution Duration**.

This alarm is cleared when **hit number** is **1** and the DNS resolution duration is less than or equal to the threshold. This alarm is cleared when **hit number** is not **1** and the DNS resolution duration is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12043	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- Kerberos-based secondary authentication is slow.
- The ZooKeeper service is abnormal.
- The node is faulty.

Possible Causes

- The node is configured with the DNS client.
- The node is equipped with the DNS server and the DNS server is started.

Procedure

Check whether the node is configured with the DNS client.

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 View the alarm details to obtain the value of the **HostName** field in **Location**.

Step 3 Use PuTTY to log in to the node for which the alarm is generated as user **root**.

Step 4 Run the **cat /etc/resolv.conf** command to check whether the DNS client is installed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If information similar to the following is displayed, the DNS client is installed and started:

```
nameserver 10.2.3.4
nameserver 10.2.3.4
```

- If yes, go to **Step 5**.
- If no, go to **Step 7**.

Step 5 Run the **vi /etc/resolv.conf** command to comment out the following content using the number signs (#) and save the file:

```
# nameserver 10.2.3.4
# nameserver 10.2.3.4
```

Step 6 Check whether this alarm is cleared after 5 minutes.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check whether the node is equipped with the DNS server and the DNS server is started.

Step 7 Run the **service named status** command to check whether the DNS service is installed on the node.

If information similar to the following is displayed, the DNS server is installed and started:

```
Checking for nameserver BIND
version: 9.6-ESV-R7-P4
CPUs found: 8
worker threads: 8
number of zones: 17
debug level: 0
xfers running: 0
xfers deferred: 0
soa queries in progress: 0
query logging is ON
recursive clients: 4/0/1000
tcp clients: 0/100
server is up and running
```

- If yes, go to **Step 8**.
- If no, go to **Step 10**.

Step 8 Run the **service named stop** command to stop the DNS server.

Step 9 Check whether this alarm is cleared after 5 minutes.

- If yes, no further action is required.
- If no, go to **Step 10**.

Step 10 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.474 ALM-12045 Read Packet Dropped Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the read packet dropped rate every 30 seconds. This alarm is generated when the read packet dropped rate exceeds the threshold (the default threshold is 0.5%) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Reading** > **Network Read Packet Rate Information** > **Read Packet Dropped Rate**.

This alarm is cleared when **hit number** is 1 and the read packet dropped rate is less than or equal to the threshold. This alarm is cleared when **hit number** is greater than 1 and the read packet dropped rate is less than or equal to 90% of the threshold.

The alarm detection is disabled by default. If you want to enable this function, check whether this function can be enabled based on Checking System Environments.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12045	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The service performance deteriorates or some services time out.

Risk warning: In SUSE kernel 3.0 or later or Red Hat 7.2, the system kernel modifies the mechanism for counting the number of dropped read packets. In this

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

case, this alarm may be generated even if the network is running properly, but services are not affected. You are advised to check the system environment first.

Possible Causes

- An OS exception occurs.
- The NICs are bonded in active/standby mode.
- The alarm threshold is improperly configured.
- The network environment is abnormal.

Procedure

View the network packet dropped rate.

Step 1 Use PuTTY to log in to any non-alarm node in the cluster as user **omm** and run the **ping** *IP address of the node for which the alarm is generated* **-c 100** command to check whether packet drop occurs on the network.

```
# ping 10.10.10.12 -c 5
PING 10.10.10.12 (10.10.10.12) 56(84) bytes of data.
64 bytes from 10.10.10.11: icmp_seq=1 ttl=64 time=0.033 ms
64 bytes from 10.10.10.11: icmp_seq=2 ttl=64 time=0.034 ms
64 bytes from 10.10.10.11: icmp_seq=3 ttl=64 time=0.021 ms
64 bytes from 10.10.10.11: icmp_seq=4 ttl=64 time=0.033 ms
64 bytes from 10.10.10.11: icmp_seq=5 ttl=64 time=0.030 ms
--- 10.10.10.12 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4001ms   rtt min/avg/max/mdev =
0.021/0.030/0.034/0.006 ms
```

NOTE

- *IP address of the node for which the alarm is generated*: Query the IP address of the node for which the alarm is generated on the node management page of the MRS cluster details page based on the value of **HostName** in the alarm location information. Check both the IP addresses of the management plane and service plane.
- **-c**: number of check times. The default value is **100**.

- If yes, go to **Step 11**.
- If no, go to **Step 2**.

Check the system environment.

Step 2 Use PuTTY to log in to the active OMS node or the node for which the alarm is generated as user **omm**.

Step 3 Run the **cat /etc/*-release** command to check the OS type.

- If the OS is EulerOS, go to **Step 4**.

```
# cat  /etc/*-release                          EulerOS release 2.0 (SP2)
EulerOS release 2.0 (SP2)
```

- If the OS is SUSE, go to **Step 5**.

```
# cat /etc/*-release
SUSE Linux Enterprise Server 11 (x86_64)
VERSION = 11
PATCHLEVEL = 3
```

- Otherwise, go to **Step 11**.

Step 4 Run the **cat /etc/euleros-release** command to check whether the OS version is EulerOS 2.2.

```
# cat/etc/euleros-release
EulerOS release 2.0 (SP2)
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, the alarm sending function cannot be enabled. Go to **Step 6**.
- If no, go to **Step 11**.

Step 5 Run the **cat /proc/version** command to check whether the SUSE kernel version is 3.0 or later.

```
# cat /proc/version
Linux version 3.0.101-63-default (geeko@buildhost) (gcc version 4.3.4 [gcc-4_3-branch revision 152973] (SUSE Linux) ) #1 SMP Tue Jun 23 16:02:31 UTC 2015 (4b89d0c)
```

- If yes, the alarm sending function cannot be enabled. Go to **Step 6**.
- If no, go to **Step 11**.

Step 6 Log in to MRS Manager and choose **System** > **Configuration** > **Threshold Configuration**.

Step 7 In the navigation pane of the **Threshold Configuration** page, choose **Network Reading** > **Network Read Packet Rate Information** > **Read Packet Dropped Rate**. In the right pane, check whether **Send Alarm** is selected.

- If yes, the alarm sending function is enabled. Go to **Step 8**.
- If no, the alarm sending function is disabled. Go to **Step 10**.

Step 8 In the right pane, deselect **Send Alarm** to shield alarm "Network Read Packet Dropped Rate Exceeds the Threshold."

Step 9 Go to the MRS cluster details page and choose **Alarms**.

Step 10 Search for alarm 12045 and manually clear the alarms that are not automatically cleared. No further action is required.

NOTE

The ID of alarm Network Read Packet Dropped Rate Exceeds the Threshold is 12045.

Check whether the NICs are bonded in active/standby mode.

Step 11 Use PuTTY to log in to the node for which the alarm is generated as user **omm** and run the **ls -l /proc/net/bonding** command to check whether the **/proc/net/bonding** directory exists on the node.

- If yes, as shown in the following figure, the bond mode is configured for the node. Go to **Step 12**.

```
# ls -l /proc/net/bonding/
total 0
-r--r--r-- 1 root root 0 Oct 11 17:35 bond0
```

- If no, the bond mode is not configured for the node. Go to **Step 14**.

```
# ls -l /proc/net/bonding/
ls: cannot access /proc/net/bonding/: No such file or directory
```

Step 12 Run the **cat /proc/net/bonding/bond0** command to check whether the value of **Bonding Mode** in the configuration file is **fault-tolerance**.

NOTE

In the preceding command, **bond0** is the name of the bond configuration file. Use the file name obtained in **Step 11**.

```
# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v3.7.1 (April 27, 2011)

Bonding Mode: fault-tolerance (active-backup)
Primary Slave: eth1 (primary_reselect always)
Currently Active Slave: eth1
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 0
Down Delay (ms): 0

Slave Interface: eth0
MII Status: up
Speed: 1000 Mbps
Duplex: full
Link Failure Count: 1
Slave queue ID: 0

Slave Interface: eth1
MII Status: up
Speed: 1000 Mbps
Duplex: full
Link Failure Count: 1
Slave queue ID: 0
```

- If yes, the NICs are bonded in active/standby mode. Go to **Step 13**.
- If no, go to **Step 14**.

Step 13 Check whether the NIC specified by **NetworkCardName** in the alarm details is the standby NIC.

- If yes, the alarm of the standby NIC cannot be automatically cleared. Manually clear the alarm on the alarm management page. No further action is required.
- If no, go to **Step 14**.

NOTE

To determine the standby NIC, check the **/proc/net/bonding/bond0** configuration file. If the NIC name corresponding to **NetworkCardName** is **Slave Interface** but not **Currently Active Slave** (the current active NIC), the NIC is the standby one.

Check whether the threshold is set properly.

Step 14 Log in to MRS Manager and check whether the threshold (configurable, 0.5% by default) is appropriate.

- If yes, go to **Step 17**.
- If no, go to **Step 15**.

Step 15 Choose **System** > **Threshold Configuration** > **Device** > **Host** > **Network Reading** > **Network Read Packet Rate Information** > **Read Packet Dropped Rate** and change the alarm threshold based on the actual service usage.

Step 16 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 17**.

Check whether the network is normal.

Step 17 Contact the system administrator to check whether the network is normal.

- If yes, rectify the network fault and go to **Step 18**.
- If no, go to **Step 19**.

Step 18 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 19**.

Step 19 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.475 ALM-12046 Write Packet Dropped Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the write packet dropped rate every 30 seconds. This alarm is generated when the write packet dropped rate exceeds the threshold (the default threshold is 0.5%) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Writing** > **Network Write Packet Rate Information** > **Write Packet Dropped Rate**.

When the **hit number** is **1**, this alarm is cleared when the network write packet dropped rate is less than or equal to the threshold. When the **hit number** is greater than **1**, this alarm is cleared when the network write packet dropped rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12046	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The service performance deteriorates or some services time out.

Possible Causes

- The alarm threshold is improperly configured.
- The network environment is abnormal.

Procedure

Check whether the threshold is set properly.

Step 1 Log in to MRS Manager and check whether the threshold (configurable, 0.5% by default) is appropriate.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **System** > **Threshold Configuration** > **Device** > **Host** > **Network Write Information** > **Network Write Packet Rate** > **Write Packet Dropped Rate** and change the alarm threshold based on the actual service usage.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network is normal.

Step 4 Contact the system administrator to check whether the network is normal.

- If yes, rectify the network fault and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Step 6 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.476 ALM-12047 Read Packet Error Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the read packet error rate every 30 seconds. This alarm is generated when the read packet error rate exceeds the threshold (the default threshold is **0.5%**) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Reading** > **Network Read Packet Rate Information** > **Read Packet Error Rate**.

If the **hit number** is **1**, this alarm is cleared when the read packet error rate is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the read packet error rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12047	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The communication is intermittently interrupted, and services time out.

Possible Causes

- The alarm threshold is improperly configured.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The network environment is abnormal.

Procedure

Check whether the threshold is set properly.

Step 1 Log in to MRS Manager and check whether the threshold (configurable, 0.5% by default) is appropriate.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **System** > **Threshold Configuration** > **Device** > **Host** > **Network Reading** > **Network Read Packet Rate Information** > **Read Packet Error Rate** and change the alarm threshold based on the actual service usage.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network is normal.

Step 4 Contact the system administrator to check whether the network is normal.

- If yes, rectify the network fault and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Step 6 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.477 ALM-12048 Write Packet Error Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the write packet error rate every 30 seconds. This alarm is generated when the write packet error rate exceeds the threshold (the default threshold is **0.5%**) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Writing** > **Network Write Packet Rate Information** > **Write Packet Error Rate**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If **hit number** is **1**, this alarm is cleared when the write packet error rate is less than or equal to the threshold. If **hit number** is greater than **1**, this alarm is cleared when the write packet error rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12048	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The communication is intermittently interrupted, and services time out.

Possible Causes

- The alarm threshold is improperly configured.
- The network environment is abnormal.

Procedure

Check whether the threshold is set properly.

Step 1 Log in to MRS Manager and check whether the threshold (configurable, 0.5% by default) is appropriate.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **System** > **Threshold Configuration** > **Device** > **Host** > **Network Writing** > **Network Write Packet Rate Information** > **Write Packet Error Rate** and change the alarm threshold based on the actual service usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network is normal.

Step 4 Contact the system administrator to check whether the network is normal.

- If yes, rectify the network fault and go to **Step 5**.
- If no, go to **Step 6**.

Step 5 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Step 6 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.478 ALM-12049 Read Throughput Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the read throughput rate every 30 seconds. This alarm is generated when the read throughput rate exceeds the threshold (the default threshold is **80%**) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Reading** > **Network Read Throughput Rate** > **Read Throughput Rate**.

If the **hit number** is **1**, this alarm is cleared when the read throughput rate is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the read throughput rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12049	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The service system runs abnormally or is unavailable.

Possible Causes

- The alarm threshold is improperly configured.
- The network port rate does not meet service requirements.

Procedure

Check whether the threshold is set properly.

Step 1 Log in to MRS Manager and check whether the threshold (configurable, 80% by default) is appropriate.

- If yes, go to **Step 2**.
- If no, go to **Step 4**.

Step 2 Choose **System** > **Threshold Configuration** > **Device** > **Host** > **Network Reading** > **Network Read Throughput Rate** > **Read Throughput Rate** to change the alarm threshold based on the actual service usage.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network port rate meets the requirements.

Step 4 In the real-time alarm list, click the alarm. In the **Alarm Details** area, obtain the IP address and network port name of the host for which the alarm is generated.

Step 5 Use PuTTY to log in to the host for which the alarm is generated as user **root**.

Step 6 Run the **ethtool** *network port name* command to check the maximum network port rate **Speed**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

In a VM environment, you may fail to obtain the network port rate by running commands. You are advised to contact the system administrator to check whether the network port rate meets the requirements.

Step 7 If the read throughput rate exceeds the threshold, contact the system administrator to increase the network port rate.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.479 ALM-12050 Write Throughput Rate Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the write throughput rate every 30 seconds. This alarm is generated when the write throughput rate exceeds the threshold (the default threshold is **80%**) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Network Writing** > **Network Write Throughput Rate** > **Write Throughput Rate**.

If the **hit number** is **1**, this alarm is cleared when the write throughput rate is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the write throughput rate is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12050	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NetworkCardName	Specifies the network port for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The service system runs abnormally or is unavailable.

Possible Causes

- The alarm threshold is improperly configured.
- The network port rate does not meet service requirements.

Procedure

Check whether the threshold is set properly.

Step 1 Log in to MRS Manager and check whether the threshold (configurable, 80% by default) is appropriate.

- If yes, go to **Step 4**.
- If no, go to **Step 2**.

Step 2 Choose **System** > **Threshold Configuration** > **Device** > **Host** > **Network Writing** > **Network Write Throughput Rate** > **Write Throughput Rate** to change the alarm threshold based on the actual service usage.

Step 3 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Check whether the network port rate meets the requirements.

Step 4 In the real-time alarm list, click the alarm. In the **Alarm Details** area, obtain the IP address and network port of the host for which the alarm is generated.

Step 5 Use PuTTY to log in to the host for which the alarm is generated as user **root**.

Step 6 Run the **ethtool** *network port name* command to check the maximum network port rate **Speed**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

In a VM environment, you may fail to obtain the network port rate by running commands. You are advised to contact the system administrator to check whether the network port rate meets the requirements.

Step 7 If the write throughput rate exceeds the threshold, contact the system administrator to increase the network port rate.

Step 8 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.480 ALM-12051 Disk Inode Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the disk inode usage every 30 seconds. This alarm is generated when the disk inode usage exceeds the threshold (the default threshold is 80%) for multiple times (the default value is 5).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Disk** > **Disk Inode Usage** > **Disk Inode Usage**.

If the **hit number** is **1**, this alarm is cleared when the disk inode usage is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the disk inode usage is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12051	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the disk partition for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Data cannot be written to the file system.

Possible Causes

- There are too many small files on the disk.
- The system is abnormal.

Procedure

There are too many small files on the disk.

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the alarm. In the **Alarm Details** area, obtain the IP address and disk partitions of the host for which the alarm is generated.

Step 3 Use PuTTY to log in to the host for which the alarm is generated as user **root**.

Step 4 Run the **df -i** *partition name* command to check the current inode usage of the disk.

Step 5 If the inode usage exceeds the threshold, manually check whether the small files in the partition can be deleted.

- If yes, delete the files and go to **Step 6**.
- If no, adjust the capacity. Then go to **Step 7**.

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

Check whether the system environment is normal.

Step 7 Contact the operating system maintenance personnel to check whether the system environment is abnormal.

- If yes, rectify the operating system fault and go to **Step 8**.
- If no, go to **Step 9**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.481 ALM-12052 Usage of Temporary TCP Ports Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the usage of temporary TCP ports every 30 seconds. This alarm is generated when the usage of temporary TCP ports exceeds the threshold (the default threshold is **80%**) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Host** > **Network Status** > **TCP Ephemeral Port Usage** > **TCP Ephemeral Port Usage**.

If the **hit number** is **1**, this alarm is cleared when the usage of temporary TCP ports is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the usage of temporary TCP ports is less than or equal to 90% of the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12052	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Services on the host fail to establish connections with the external and services are interrupted.

Possible Causes

- The temporary ports do not meet service requirements.
- The system is abnormal.

Procedure

Expand the range of temporary ports.

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the alarm. In the **Alarm Details** area, obtain the IP address of the host for which the alarm is generated.

Step 3 Use PuTTY to log in to the host for which the alarm is generated as user **omm**.

Step 4 Run the **cat /proc/sys/net/ipv4/ip_local_port_range |cut -f 1** command to obtain the start port number. Run the **cat /proc/sys/net/ipv4/ip_local_port_range |cut -f 2** command to obtain the end port number. Subtract the start port number from the end port number to obtain the total number of temporary ports. If the total number of temporary ports is less than 28,232, the random port range of the OS is too small. In this case, contact the system administrator to expand the port range.

Step 5 Run the **ss -ant 2>/dev/null | grep -v LISTEN | awk 'NR > 2 {print $4}'|cut -d ':' -f 2 | awk '$1 >"***start port number***" {print $1}' | sort -u | wc -l** command to calculate the number of used temporary ports.

Step 6 Calculate the usage of temporary ports using the following formula: Usage of temporary ports = (Number of used temporary ports/Total number of temporary ports) x 100. Check whether the usage exceeds the threshold.

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

Step 7 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 8**.

Check whether the system environment is normal.

Step 8 Run the following command to import the temporary file and view the frequently used ports in the **port_result.txt** file:

netstat -tnp > $BIGDATA_HOME/tmp/port_result.txt

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
netstat -tnp

Active Internet connections (w/o servers)

Proto Recv Send LocalAddress ForeignAddress State PID/ProgramName tcp   0   0 10-120-85-154:45433
10-120-8:25009 CLOSE_WAIT 94237/java
tcp   0   0 10-120-85-154:45434  10-120-8:25009 CLOSE_WAIT 94237/java
tcp   0   0 10-120-85-154:45435  10-120-8:25009 CLOSE_WAIT 94237/java
...
```

Step 9 Run the following command to check the processes that occupy a large number of ports:

ps -ef |grep *PID*

NOTE

- *PID* indicates the process ID of the port queried in **Step 8**.
- Run the following command to collect information about all processes in the system and check the processes that occupy a large number of ports:

  ps -ef > $BIGDATA_HOME/tmp/ps_result.txt

Step 10 Contact the system administrator to clear the processes that occupy a large number of ports. Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Step 11 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.482 ALM-12053 File Handle Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the handler usage every 30 seconds. This alarm is generated when the handle usage exceeds the threshold (the default threshold is **80%**) for multiple times (the default value is **5**).

You can change the threshold by choosing **System** > **Threshold Configuration** > **Device** > **Host** > **Host Status** > **Host File Handle Usage** > **Host File Handle Usage**.

If the **hit number** is **1**, this alarm is cleared when the file handle usage is less than or equal to the threshold. If the **hit number** is greater than **1**, this alarm is cleared when the file handle usage is less than or equal to 90% of the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12053	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

The system applications fail to open files, access networks, and perform other I/O operations. The applications are running improperly.

Possible Causes

- The number of file handles does not meet service requirements.
- The system is abnormal.

Procedure

Increase the number of file handles.

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the alarm. In the **Alarm Details** area, obtain the IP address of the host for which the alarm is generated.

Step 3 Use PuTTY to log in to the host for which the alarm is generated as user **root**.

Step 4 Run the **ulimit -n** command to check the maximum number of handles set in the system.

Step 5 If the file handle usage exceeds the threshold, contact the system administrator to increase the number of system file handles.

Step 6 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 7**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the system environment is normal.

Step 7 Contact the system administrator to check whether the OS is abnormal.

- If yes, rectify the operating system fault and go to **Step 8**.
- If no, go to **Step 9**.

Step 8 Wait 5 minutes and check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Step 9 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.483 ALM-12054 Invalid Certificate File (For MRS 2.*x* or Earlier)

Description

The system checks whether the certificate file is invalid (has expired or is not yet valid) on 23:00 every day. This alarm is generated when the certificate file is invalid.

This alarm is cleared when the status of the newly imported certificate is valid.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12054	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The system reminds users that the certificate file is invalid. If the certificate file is invalid, some functions are restricted and cannot be used properly.

Possible Causes

No certificate (HA root certificate or HA user certificate) is imported to the system, the certificate fails to be imported, or the certificate file is invalid.

Procedure

Check the alarm cause.

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the row that contains the alarm.

In the **Alarm Details** area, view the additional information about the alarm.

- If **CA Certificate** is displayed in the additional alarm information, use PuTTY to log in to the active OMS management node as user **omm** and go to **Step 3**.
- If **HA root Certificate** is displayed in the additional information, check **Location** to obtain the name of the host involved in this alarm. Then use PuTTY to log in to the host as user **omm** and go to **Step 4**.
- If **HA server Certificate** is displayed in the additional information, check **Location** to obtain the name of the host involved in this alarm. Then use PuTTY to log in to the host as user **omm** and go to **Step 5**.

Check the validity period of the certificate files in the system.

Step 3 Check whether the current system time is in the validity period of the CA certificate.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/cert/root/ca.crt** command to check the effective time and due time of the root certificate.

- If yes, go to **Step 8**.
- If no, go to **Step 6**.

Step 4 Check whether the current system time is in the validity period of the HA root certificate.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/root-ca.crt** command to check the effective time and due time of the HA root certificate.

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

Step 5 Check whether the current system time is in the validity period of the HA user certificate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/server.crt** command to check the effective time and due time of the HA user certificate.

- If yes, go to **Step 8**.
- If no, go to **Step 7**.

  The following is an example of the effective time and expiration time of a CA or HA certificate:

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            97:d5:0e:84:af:ec:34:d8
        Signature Algorithm: sha256WithRSAEncryption
        Issuer: C=CountryName, ST=State, L=Locality, O=Organization, OU=IT, CN=HADOOP.COM
        Validity
            Not Before: Dec 13 06:38:26 2016 GMT            // Effective time
            Not After : Dec 11 06:38:26 2026 GMT            // Expiration time
```

Import certificate files.

Step 6 Import a new CA certificate file.

Contact O&M personnel to apply for or generate a new CA certificate file and import it. Manually clear the alarm and check whether this alarm is generated again during periodic check.

- If yes, go to **Step 8**.
- If no, no further action is required.

Step 7 Import a new HA certificate file.

Apply for or generate a new HA certificate file and import it. Manually clear the alarm and check whether this alarm is generated again during periodic check.

- If yes, go to **Step 8**.
- If no, no further action is required.

Step 8 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.484 ALM-12055 Certificate File Is About to Expire (For MRS 2.*x* or Earlier)

Description

The system checks the certificate file on 23:00 every day. This alarm is generated if the certificate file is about to expire with a validity period less than days set in the alarm threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is generated if the status of the newly imported certificate is valid.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12055	Minor	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The system reminds users that the license is about to expire. If the license expires, some functions are restricted and cannot be used properly.

Possible Causes

The remaining validity period of the CA certificate, HA root certificate, or HA user certificate is smaller than the alarm threshold.

Procedure

Check the alarm cause.

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the real-time alarm list, click the row that contains the alarm.

In the **Alarm Details** area, view the additional information about the alarm.

- If **CA Certificate** is displayed in the additional alarm information, use PuTTY to log in to the active OMS management node as user **omm** and go to **Step 3**.
- If **HA root Certificate** is displayed in the additional information, check **Location** to obtain the name of the host involved in this alarm. Then use PuTTY to log in to the host as user **omm** and go to **Step 4**.
- If **HA server Certificate** is displayed in the additional information, check **Location** to obtain the name of the host involved in this alarm. Then use PuTTY to log in to the host as user **omm** and go to **Step 5**.

Check the validity period of the certificate files in the system.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the remaining validity period of the CA certificate is smaller than the alarm threshold.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/cert/root/ca.crt** command to check the effective time and due time of the root certificate.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 4 Check whether the remaining validity period of the HA root certificate is smaller than the alarm threshold.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/root-ca.crt** command to check the effective time and due time of the HA root certificate.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

Step 5 Check whether the remaining validity period of the HA user certificate is smaller than the alarm threshold.

Run the **openssl x509 -noout -text -in ${CONTROLLER_HOME}/security/certHA/server.crt** command to check the effective time and due time of the HA user certificate.

- If yes, go to **Step 7**.
- If no, go to **Step 8**.

  The following is an example of the effective time and expiration time of a CA or HA certificate:

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            97:d5:0e:84:af:ec:34:d8
        Signature Algorithm: sha256WithRSAEncryption
        Issuer: C=CountryName, ST=State, L=Locality, O=Organization, OU=IT, CN=HADOOP.COM
        Validity
            Not Before: Dec 13 06:38:26 2016 GMT            // Effective time
            Not After : Dec 11 06:38:26 2026 GMT            // Expiration time
```

Import certificate files.

Step 6 Import a new CA certificate file.

Contact O&M personnel to apply for or generate a new CA certificate file and import it. Manually clear the alarm and check whether this alarm is generated again during periodic check.

- If yes, go to **Step 8**.
- If no, no further action is required.

Step 7 Import a new HA certificate file.

Apply for or generate a new HA certificate file and import it. Manually clear the alarm and check whether this alarm is generated again during periodic check.

- If yes, go to **Step 8**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, no further action is required.

Step 8 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.485 ALM-12180 Disk Card I/O (For MRS 2.*x* or Earlier)

Description

For MRS 2.*x* or earlier:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - The system collects data every 3 seconds, and detects that the **svctm** value exceeds 6s for 10 consecutive periods within 30 seconds.
  - The system collects data every 3 seconds, and detects that the **avgqu-sz** value is greater than 0, the IOPS or bandwidth is 0, and the **ioutil** value is greater than **99%** for 10 consecutive periods within 30 seconds.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - The system collects data every 3 seconds, and detects that the **svctm** value exceeds 2s for 10 consecutive periods within 30 seconds.
  - The system collects data every 3 seconds, and detects that the **avgqu-sz** value is greater than 0, the IOPS or bandwidth is 0, and the **ioutil** value is greater than **99%** for 10 consecutive periods within 30 seconds.

This alarm is automatically cleared when none of the conditions are met for 90 seconds.

For MRS 1.9.3.10 or later:

- For HDDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 6 seconds within 30 seconds in at least seven collection periods.
  - The system collects data every 3 seconds by default, and within 30 seconds, at least 10 collection cycles have a disk queue depth (**avgqu-sz**) > 0, IOPS = 0, or bandwidth = 0, and **ioutil** > 99%.
  - By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 1000 ms within 300 seconds.
- For SSDs, the alarm is triggered when any of the following conditions is met:
  - By default, the system collects data every 3 seconds. The svctm latency reaches 3 seconds within 30 seconds in at least seven collection periods.
  - By default, the system collects data every 3 seconds. Disk queue depth (**avgqu-sz**) > 0 and IOPS = 0, or bandwidth = 0 and **ioutil** > 99% in at least 10 collection periods within 30 seconds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- By default, the system collects data every 3 seconds. At least 50% of detected svctm take no less than 500 ms within 300 seconds.

The collection period is 3 seconds, and the detection period is 30 or 300 seconds. This alarm is automatically cleared when none of the conditions are met for three consecutive detection periods (30 or 300 seconds).

NOTE

For details about how to obtain related parameters, see **Related Information**.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12180	Major	Yes

Parameters

Parameter	Description
Source	Specifies the cluster or system for which the alarm is generated.
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
DiskName	Specifies the disk for which the alarm is generated.

Impact on the System

A continuously high I/O usage may adversely affect service operations and result in service loss.

Possible Causes

The disk is aged.

Procedure

Replace the disk.

Step 1 Log in to FusionInsight Manager and choose **O&M** > **Alarm** > **Alarms**.

Step 2 View the detailed information about the alarm. Check the values of **HostName** and **DiskName** in the location information to obtain the information about the faulty disk for which the alarm is reported.

Step 3 Replace the faulty disk.

Step 4 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, perform **Step 5**.

Collect the fault information.

Step 5 On MRS Manager, choose **System** > **Export Log**.

Step 6 Contact O&M engineers and send the collected logs.

----End

Alarm Clearing

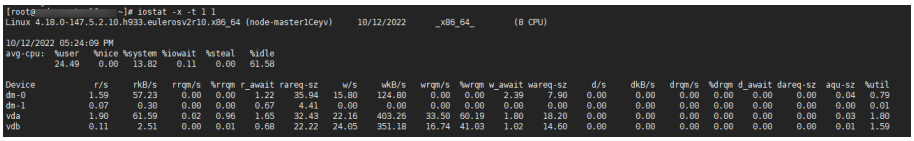
This alarm is automatically cleared after the fault is rectified.

Related Information

To obtain the related parameters, perform the following steps:

- Run the following command in the OS to collect data:

  iostat -x -t 1 1

  The command parameters are as follows:

  avgqu-sz indicates the disk queue depth.

  The sum of **r/s** and **w/s** is the IOPS.

  The sum of **rkB/s** and **wkB/s** is the bandwidth.

  %util is the value of **ioutil**.

- The value of **svctm** is calculated as follows:

  svctm = (tot_ticks_new - tot_ticks_old) / (rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old)

  For MRS 2.*x* or earlier:

  If **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old** is **0**, then **svctm** is **0**.

  For MRS 1.9.3.10 or later:

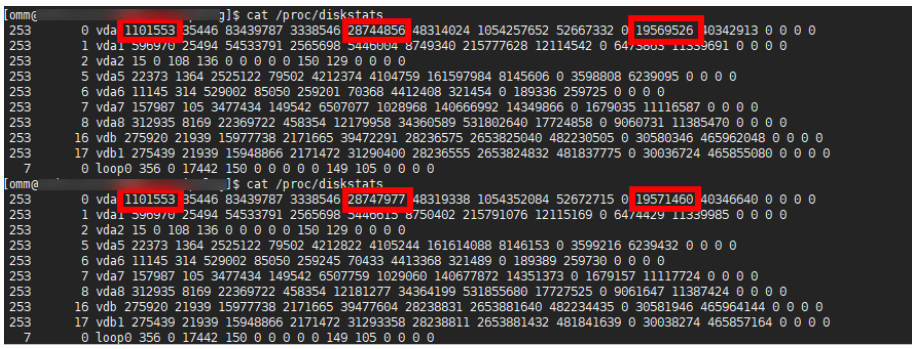
  When the detection period is 30 seconds, if **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, then **svctm = 0**.

  When the detection period is 300 seconds and **rd_ios_new + wr_ios_new - rd_ios_old - wr_ios_old = 0**, if **tot_ticks_new - tot_ticks_old = 0**, then **svctm = 0**; otherwise, the value of **svctm** is infinite.

  The parameters in the preceding expression can be obtained as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Obtain the parameter values from the data collected via the **cat /proc/diskstats** command run by the system every 3 seconds. The following shows an example.



In the data collected for the first time, the number in the fourth column is the value of **rd_ios_old**, the number in the eighth column is the value of **wr_ios_old**, and the number in the thirteenth column is the value of **tot_ticks_old**.

In the data collected for the second time, the number in the fourth column is the value of **rd_ios_new**, the number in the eighth column is the value of **wr_ios_new**, and the number in the thirteenth column is the value of **tot_ticks_new**.

In this case, the value of **svctm** is as follows:

(19571460 - 19569526) / (1101553 + 28747977 - 1101553 - 28744856) = 0.6197

9.11.486 ALM-12357 Failed to Export Audit Logs to OBS (For MRS 2.*x* or Earlier)

Description

If audit log transfer to OBS has been configured on MRS Manager, the system periodically transfers audit logs to OBS. This alarm is generated when the system cannot access OBS.

This alarm is cleared after the system exports audit logs to the OBS successfully.

Attribute

Alarm ID	Alarm Severity	Auto Clear
12357	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The local system saves a maximum of seven compressed service audit log files. If this alarm persists, local service audit logs may be lost.

The local system saves a maximum of 50 management audit log files (each file contains 100,000 records). If this alarm persists, local management audit logs may be lost.

Possible Causes

- Connection to the OBS server fails.
- The specified OBS file system does not exist.
- The user AK/SK information is invalid.
- The local OBS configuration cannot be obtained.

Procedure

Step 1 Log in to the OBS server and check whether the OBS server can be properly accessed.

- If yes, go to **Step 3**.
- If no, go to **Step 2**.

Step 2 Contact the maintenance personnel to repair OBS. Then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 On MRS Manager, choose **System** > **Export Audit Log**. Check whether the AK/SK information, file system name, and path are correct.

- If yes, go to **Step 5**.
- If no, go to **Step 4**.

Step 4 Correct the information. Then check whether the alarm is cleared when the export task is executed again.

NOTE

To check alarm clearance quickly, you can set the start time of audit log collection to 10 or 30 minutes later than the current time. After checking the result, restore the original start time.

- If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

9.11.487 ALM-13000 ZooKeeper Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the ZooKeeper service status every 30 seconds. This alarm is generated when the ZooKeeper service is unavailable.

This alarm is cleared when the ZooKeeper service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13000	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

ZooKeeper fails to provide coordination services for upper-layer components and the components depending on ZooKeeper may not run properly.

Possible Causes

- The ZooKeeper instance is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The disk capacity is insufficient.
- The network is faulty.
- The DNS is installed on the ZooKeeper node.

Procedure

Check the ZooKeeper service instance status.

Step 1 On the MRS cluster details page, choose **Components** > **ZooKeeper** > **quorumpeer**.

Step 2 Check whether the ZooKeeper instances are normal.

- If yes, go to **Step 6**.
- If no, go to **Step 3**.

Step 3 Select instances whose status is not good and choose **More** > **Restart Instance**.

Step 4 Check whether the instance status is good after restart.

- If yes, go to **Step 5**.
- If no, go to **Step 19**.

Step 5 On the **Alarms** tab page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 6**.

Check disk status.

Step 6 On the MRS cluster details page, choose **Components** > **ZooKeeper** > **quorumpeer**, and check the host information of each node housing the ZooKeeper instance.

Step 7 On the MRS cluster details page, click the **Nodes** tab and expand a node group.

Step 8 In the **Disk Usage** column, check whether the disk space of each node housing ZooKeeper instances is insufficient (disk usage exceeds 80%).

- If yes, go to **Step 9**.
- If no, go to **Step 11**.

Step 9 Expand the disk capacity. For details, see **ALM-12017 Insufficient Disk Capacity (For MRS 2.x or Earlier)**.

Step 10 On the **Alarms** tab page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 11**.

Check network communication status.

Step 11 On the Linux node housing the ZooKeeper instance, run the **ping** command to check whether the host names of other nodes housing the ZooKeeper instances can be pinged successfully.

- If yes, go to **Step 15**.
- If no, go to **Step 12**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 12 Modify the IP addresses in **/etc/hosts** and add the mapping between host names and IP addresses.

Step 13 Run the **ping** command again to check whether the host names of other nodes housing the ZooKeeper instances can be pinged successfully.

- If yes, go to **Step 14**.
- If no, go to **Step 19**.

Step 14 On the **Alarms** tab page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 15**.

Check the DNS.

Step 15 Check whether the DNS is installed on the node housing the ZooKeeper instance. On the Linux node housing the ZooKeeper instance, run the **cat /etc/resolv.conf** command to check whether the file is empty.

- If yes, go to **Step 16**.
- If no, go to **Step 19**.

Step 16 Run the **service named status** command to check whether the DNS is started.

- If yes, go to **Step 17**.
- If no, go to **Step 19**.

Step 17 Run the **service named stop** command to stop the DNS service. If "Shutting down name server BIND waiting for named to shut down (28s)" is displayed, the DNS service is stopped successfully. Comment out the content (if any) in **/etc/resolv.conf**.

Step 18 On the **Alarms** tab page, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 19**.

Step 19 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.488 ALM-13001 Available ZooKeeper Connections Are Insufficient (For MRS 2.*x* or Earlier)

Description

The system checks ZooKeeper connections every 30 seconds. This alarm is generated when the system detects that the number of used ZooKeeper instance connections exceeds the threshold (80% of the maximum connections).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the number of used ZooKeeper instance connections is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13001	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Available ZooKeeper connections are insufficient. When the connection usage reaches 100%, external connections cannot be handled.

Possible Causes

The number of connections to the ZooKeeper node exceeds the threshold. Connection leakage occurs on some connection processes, or the maximum number of connections does not meet the requirement of the actual scenario.

Procedure

Step 1 Check the connection status.

1. On the MRS cluster details page, choose **Alarms** > **ALM-13001 Available ZooKeeper Connections Are Insufficient** > **Location**. Check the IP address of the node for which the alarm is generated.
2. Obtain the PID of the ZooKeeper process. Log in to the node for which this alarm is generated and run the **pgrep -f proc_zookeeper** command.
3. Check whether the PID can be successfully obtained.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Obtain all the IP addresses connected to the ZooKeeper instance and the number of connections and check 10 IP addresses with top connections. Run the **lsof -i|grep $pid | awk '{print $9}' | cut -d : -f 2 | cut -d \> -f 2 | awk '{a[$1]++} END {for(i in a){print i,a[i] | "sort -r -g -k 2"}}' | head -10** command based on the obtained PID value. (**$pid** is the PID obtained in the preceding step.)
5. Check whether the node IP addresses and the number of connections are successfully obtained.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 2**.
6. Obtain the ID of the port connected to the process. Run the **lsof -i|grep $pid | awk '{print $9}'|cut -d \> -f 2 |grep $IP| cut -d : -f 2** command based on the obtained PID and IP address. (**$pid** and **$IP** are the PID and IP address obtained in the preceding step.)
7. Check whether the port ID is successfully obtained.
   - If yes, go to **Step 1.8**.
   - If no, go to **Step 2**.
8. Obtain the ID of the connected process. Log in to each IP address and run the following command based on the obtained port ID: **lsof -i|grep $port**. (**$port** is the port ID obtained in the preceding step.)
9. Check whether the process ID is successfully obtained.
   - If yes, go to **Step 1.10**.
   - If no, go to **Step 2**.
10. Check whether connection leakage occurs on the process based on the obtained process ID.
    - If yes, go to **Step 1.11**.
    - If no, go to **Step 1.12**.
11. Close the process where connection leakage occurs and check whether the alarm is cleared.
    - If yes, no further action is required.
    - If no, go to **Step 1.12**.
12. On the MRS cluster details page, choose **Components** > **ZooKeeper** > **Service Configuration**. Set **Type** to **All**, choose **quorumpeer** > **Performance**, and change the value of **maxCnxns** to **20000** or more.
13. Check whether the alarm is cleared.
    - If yes, no further action is required.
    - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.489 ALM-13002 ZooKeeper Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the ZooKeeper service status every 30 seconds. The alarm is generated when the memory usage of a ZooKeeper instance exceeds the threshold (80% of the maximum memory).

The alarm is cleared when the memory usage is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
13002	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

If the available ZooKeeper memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The memory usage of the ZooKeeper instance is overused or the memory is inappropriately allocated.

Procedure

Step 1 Check the memory usage.

1. On the MRS cluster details page, choose **Alarms** > **ALM-13002 ZooKeeper Memory Usage Exceeds the Threshold** > **Location**. Check the IP address of the instance for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. On the MRS cluster details page, choose **Components** > **ZooKeeper** > **Instances** > **quorumpeer** (IP address of the instance for which the alarm is generated) > **Customize** > **ZooKeeper Heap And Direct Buffer Resource**. Check the heap memory usage.
3. Check whether the used heap memory of ZooKeeper reaches 80% of the maximum heap memory specified for ZooKeeper.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 1.6**.
4. On MRS Manager, choose **Services** > **ZooKeeper** > **Configuration** > **All** > **quorumpeer** > **System**. Increase the value of **-Xmx** in **GC_OPTS** as required.
5. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 1.6**.
6. On the MRS cluster details page, choose **Components** > **ZooKeeper** > **Instances** > **quorumpeer** (IP address of the instance for which the alarm is generated) > **Customize** > **ZooKeeper Heap And Direct Buffer Resource**. Check the direct buffer memory usage.
7. Check whether the used direct buffer memory of ZooKeeper reaches 80% of the maximum direct buffer memory specified for ZooKeeper.
   - If yes, go to **Step 1.8**.
   - If no, go to **Step 2**.
8. On the MRS cluster details page, choose **Components** > **ZooKeeper** > **Service Configuration**. Set **Type** to **All** and choose **quorumpeer** > **System**.

   Increase the value of **-XX:MaxDirectMemorySize** in **GC_OPTS** as required.
9. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.490 ALM-14000 HDFS Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the service status of NameService every 30 seconds. This alarm is generated when the system considers that the HDFS service is unavailable because all the NameService services are abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when at least one NameService service is normal and the system considers that the HDFS service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14000	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

HDFS fails to provide services for HDFS service-based upper-layer components, such as HBase and MapReduce. As a result, users cannot read or write files.

Possible Causes

- ZooKeeper is abnormal.
- All NameService services are abnormal.

Procedure

Step 1 Check the ZooKeeper status.

1. Go to the MRS cluster details page. On the **Components** tab page, check whether the health status of the ZooKeeper service is **Good**.
   - If yes, go to **Step 1.2**.
   - If no, go to **Step 2.1**.
2. Rectify the health status of the ZooKeeper service. For details, see **ALM-13000 ZooKeeper Service Unavailable (For MRS 2.x or Earlier)**. Then check whether the health status of the ZooKeeper service is **Good**.
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 3**.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 2.1**.

Step 2 Handle the NameService service exception alarm.

1. Go to the MRS cluster details page. On the **Alarms** page, check whether all NameService services have abnormal alarms.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 3**.
2. Handle the abnormal NameService services following the instructions in **ALM-14010 NameService Is Abnormal (For MRS 2.x or Earlier)** and check whether each NameService service exception alarm is cleared.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.491 ALM-14001 HDFS Disk Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the disk usage of the HDFS cluster every 30 seconds and compares the actual disk usage with the threshold. The HDFS cluster disk usage indicator has a default threshold. This alarm is generated when the HDFS disk usage exceeds the threshold.

This alarm is cleared when the disk usage of the HDFS cluster is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14001	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NSName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

The performance of writing data to HDFS is affected.

Possible Causes

The disk space configured for the HDFS cluster is insufficient.

Procedure

Step 1 Check the disk capacity and delete unnecessary files.

1. On the MRS cluster details page, choose **Components** > **HDFS**. The **Service Status** page is displayed.
2. In the **Charts** area, view the value of the monitoring indicator **Percentage of HDFS Capacity** to check whether the HDFS disk usage exceeds the threshold (80% by default).
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 3**.
3. Use the client on the cluster node and run the **hdfs dfsadmin -report** command to check whether the value of **DFS Used%** is less than 100% minus the threshold.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 3**.
4. Use the client on the cluster node and run the **hdfs dfs -rm -r** *file or directory path* command to delete unnecessary files.
5. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 Expand the system.

1. Expand the disk capacity.
2. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.492 ALM-14002 DataNode Disk Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the DataNode disk usage every 30 seconds and compares the actual disk usage with the threshold. The **Percentage of DataNode Capacity** indicator has a default threshold. This alarm is generated when the value of the **Percentage of DataNode Capacity** indicator exceeds the threshold.

This alarm is cleared when the value of the **Percentage of DataNode Capacity** indicator is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14002	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Insufficient disk space will impact read/write to HDFS.

Possible Causes

- The disk space configured for the HDFS cluster is insufficient.
- Data skew occurs among DataNodes.

Procedure

Step 1 Check the cluster disk capacity.

1. Go to the MRS cluster details page. On the **Alarms** page, check whether the ALM-14001 HDFS Disk Usage Exceeds the Threshold alarm exists.
   - If yes, go to **Step 1.2**.
   - If no, go to **Step 2.1**.
2. Handle the alarm by following the instructions in ALM-14001 HDFS Disk Usage Exceeds the Threshold and check whether the alarm is cleared.
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 3**.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the balance status of DataNodes.

1. Use the client on the cluster node, run the **hdfs dfsadmin -report** command to view the value of **DFS Used%** on the DataNode for which the alarm is generated, and compare the value with those on other DataNodes. Check whether the difference between the values is larger than 10.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 3**.
2. If data skew occurs, use the client on the cluster node and run the **hdfs balancer -threshold 10** command.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.493 ALM-14003 Number of Lost HDFS Blocks Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the number of lost blocks every 30 seconds and compares the number of lost blocks with the threshold. The lost blocks indicator has a default threshold. This alarm is generated when the number of lost blocks exceeds the threshold.

This alarm is cleared when the number of lost blocks is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14003	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NSName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Data stored in HDFS is lost. HDFS may enter the safe mode and cannot provide write services. Lost block data cannot be restored.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The DataNode instance is abnormal.
- Data is deleted.

Procedure

Step 1 Check the DataNode instance.

1. On the MRS cluster details page, choose **Components** > **HDFS** > **Instances**.
2. Check whether the status of all DataNode instances is **Good**.
   - If yes, go to **Step 3**.
   - If no, go to **Step 1.3**.
3. Restart the DataNode instance and check whether the restart is successful.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 2.1**.

Step 2 Delete the damaged file.

1. Use the client on the cluster node. Run the **hdfs fsck / -delete** command to delete the lost file. Then rewrite the file and recover the data.
2. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.494 ALM-14004 Number of Damaged HDFS Blocks Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the number of damaged blocks every 30 seconds and compares the number of damaged blocks with the threshold. The damaged blocks indicator has a default threshold. This alarm is generated when the number of damaged blocks exceeds the threshold.

This alarm is cleared when the number of damaged blocks is less than or equal to the threshold. You are advised to run the **hdfs fsck /** command to check whether any file is completely damaged.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14004	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NSName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Data is damaged and HDFS fails to read files.

Possible Causes

- The DataNode instance is abnormal.
- Data verification information is damaged.

Procedure

Step 1 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.495 ALM-14006 Number of HDFS Files Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system periodically checks the number of HDFS files every 30 seconds and compares the number of HDFS files with the threshold. This alarm is generated when the system detects that the number of HDFS files exceeds the threshold.

This alarm is cleared when the number of HDFS files is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14006	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NSName	Specifies the NameService service for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Disk storage space is insufficient, which may result in data import failure. The performance of the HDFS system is affected.

Possible Causes

The number of HDFS files exceeds the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Check whether unnecessary files exist in the system.

1. Use the client on the cluster node and run the **hdfs dfs -ls** *file or directory path* command to check whether the file or directory can be deleted.
   - If yes, go to **Step 1.2**.
   - If no, go to **Step 2.1**.
2. Run the **hdfs dfs -rm -r** *file or directory path* command. Delete unnecessary files, wait 5 minutes, and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the number of files in the system.

1. On MRS Manager, choose **System** > **Threshold Configuration**.
2. In the navigation tree on the left, choose **Services** > **HDFS** > **HDFS File** > **Total Number of Files**.
3. In the right pane, modify the threshold in the rule based on the number of current HDFS files.

   To check the number of HDFS files, choose **Services** > **HDFS**, click **Customize** in the **Real-Time Statistics** area on the right, and select the **HDFS File** monitoring item.
4. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.496 ALM-14007 HDFS NameNode Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the HDFS NameNode memory usage every 30 seconds and compares the actual memory usage with the threshold. The HDFS NameNode memory usage has a default threshold. This alarm is generated when the HDFS NameNode memory usage exceeds the threshold.

This alarm is cleared when the HDFS NameNode memory usage is less than or equal to the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14007	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

If the memory usage of the HDFS NameNode is too high, data read/write performance of HDFS will be affected.

Possible Causes

The HDFS NameNode memory is insufficient.

Procedure

Step 1 Delete unnecessary files.

1. Use the client on the cluster node and run the **hdfs dfs -rm -r** *file or directory path* command to delete unnecessary files.
2. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.497 ALM-14008 HDFS DataNode Memory Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the HDFS DataNode memory usage every 30 seconds and compares the actual memory usage with the threshold. The HDFS DataNode memory usage has a default threshold. This alarm is generated when the HDFS DataNode memory usage exceeds the threshold.

This alarm is cleared when the HDFS DataNode memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14007	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

The HDFS DataNode memory usage is too high, which affects the data read/write performance of the HDFS.

Possible Causes

The HDFS DataNode memory is insufficient.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Delete unnecessary files.

1. Use the client on the cluster node and run the **hdfs dfs -rm -r** *file or directory path* command to delete unnecessary files.
2. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.498 ALM-14009 Number of Faulty DataNodes Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system periodically checks the number of faulty DataNodes in the HDFS cluster every 30 seconds, and compares the number with the threshold. The number of faulty DataNodes has a default threshold. This alarm is generated when the number of faulty DataNodes in the HDFS cluster exceeds the threshold.

This alarm is cleared when the number of faulty DataNodes in the HDFS cluster is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14009	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Faulty DataNodes cannot provide HDFS services.

Possible Causes

- DataNodes are faulty or overloaded.
- The network between the NameNode and the DataNode is disconnected or busy.
- NameNodes are overloaded.

Procedure

Step 1 Check whether DataNodes are faulty.

1. Use the client on the cluster node to run the **hdfs dfsadmin -report** command to check whether DataNodes are faulty.
   - If yes, go to **Step 1.2**.
   - If no, go to **Step 2.1**.
2. On the MRS cluster details page, choose **Components** > **HDFS** > **Instances** to check whether the DataNode is stopped.
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 2.1**.
3. Select the DataNode instance, and choose **More** > **Restart Instance** to restart it. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the status of the network between the NameNode and the DataNode.

1. Log in to the service IP address of the node where the faulty DataNode is located, and run the **ping** *IP address of the NameNode* command to check whether the network between the DataNode and the NameNode is abnormal.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 3.1**.
2. Rectify the network fault. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 3.1**.

Step 3 Check whether the DataNode is overloaded.

1. On the MRS cluster details page, click **Alarms** and check whether the alarm ALM-14008 HDFS DataNode Memory Usage Exceeds the Threshold exists.
   - If yes, go to **Step 3.2**.
   - If no, go to **Step 4.1**.
2. Follow procedures in **ALM-14008 HDFS DataNode Memory Usage Exceeds the Threshold (For MRS 2.x or Earlier)** to handle the alarm and check whether the alarm is cleared.
   - If yes, go to **Step 3.3**.
   - If no, go to **Step 4.1**.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4.1**.

Step 4 Check whether the NameNode is overloaded.

1. On the MRS cluster details page, click **Alarms** and check whether the alarm ALM-14007 HDFS NameNode Memory Usage Exceeds the Threshold exists.
   - If yes, go to **Step 4.2**.
   - If no, go to **Step 5**.
2. Follow procedures in **ALM-14007 HDFS NameNode Memory Usage Exceeds the Threshold (For MRS 2.x or Earlier)** to handle the alarm and check whether the alarm is cleared.
   - If yes, go to **Step 4.3**.
   - If no, go to **Step 5**.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.499 ALM-14010 NameService Is Abnormal (For MRS 2.*x* or Earlier)

Description

The system checks the NameService service status every 180 seconds. This alarm is generated when the NameService service is unavailable.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

This alarm is cleared when the NameService service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14010	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
NSName	Specifies the NameService service for which the alarm is generated.

Impact on the System

HDFS fails to provide services for upper-layer components based on the NameService service, such as HBase and MapReduce. As a result, users cannot read or write files.

Possible Causes

- The JournalNode is faulty.
- The DataNode is faulty.
- The disk capacity is insufficient.
- The NameNode enters safe mode.

Procedure

Step 1 Check the status of the JournalNode instance.

1. On the MRS Manager home page, click **Components**.
2. Click **HDFS**.
3. Click **Instance**.
4. Check whether the **Health Status** of the JournalNode is **Good**.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.5**.
5. Select the faulty JournalNode, and choose **More** > **Restart Instance**. Check whether the JournalNode successfully restarts.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 1.6**.
- If no, go to **Step 5**.

6. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the status of the DataNode instance.

1. On the MRS cluster details page, click **Components**.
2. Click **HDFS**.
3. In **Operation and Health Summary**, check whether the **Health Status** of all DataNodes is **Good**.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 2.4**.
4. Click **Instances**. On the DataNode management page, select the faulty DataNode, and choose **More** > **Restart Instance**. Check whether the DataNode successfully restarts.
   - If yes, go to **Step 2.5**.
   - If no, go to **Step 3.1**.
5. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4.1**.

Step 3 Check the disk status.

1. On the MRS cluster details page, click the **Nodes** tab and expand a node group.
2. In the **Disk Usage** column, check whether disk space is insufficient.
   - If yes, go to **Step 3.3**.
   - If no, go to **Step 4.1**.
3. Expand the disk capacity.
4. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4.1**.

Step 4 Check whether NameNode is in the safe mode.

1. Use the client on the cluster node, and run the **hdfs dfsadmin -safemode get** command to check whether **Safe mode is ON** is displayed.

   Information behind **Safe mode is ON** is alarm information and is displayed based actual conditions.
   - If yes, go to **Step 4.2**.
   - If no, go to **Step 5**.
2. Use the client on the cluster node and run the **hdfs dfsadmin -safemode leave** command.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.500 ALM-14011 HDFS DataNode Data Directory Is Not Configured Properly (For MRS 2.*x* or Earlier)

Description

The DataNode parameter **dfs.datanode.data.dir** specifies the DataNode data directory. This alarm is generated in any of the following scenarios:

- A configured data directory cannot be created.
- A data directory uses the same disk as other critical directories in the system.
- Multiple directories use the same disk.

This alarm is cleared when the DataNode data directory is configured properly and this DataNode is restarted.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14011	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the DataNode data directory is mounted on critical directories such as the root directory, the disk space of the root directory will be used up after running for a long time. This causes a system fault.

If the DataNode data directory is not configured properly, HDFS performance will deteriorate.

Possible Causes

- The DataNode data directory fails to be created.
- The DataNode data directory uses the same disk as critical directories, such as **/** or **/boot**.
- Multiple directories in the DataNode data directory use the same disk.

Procedure

Step 1 Check the alarm cause and information about the DataNode for which the alarm is generated.

1. On the MRS cluster details page, click **Alarms**. In the alarm list, click the alarm.
2. In the **Alarm Details** area, view **Alarm Cause** to obtain the cause of the alarm. In **HostName** of **Location**, obtain the host name of the DataNode for which the alarm is generated.

Step 2 Delete directories that do not comply with the disk plan from the DataNode data directory.

1. Choose **Components** > **HDFS** > **Instances**. In the instance list, click the DataNode instance on the node for which the alarm is generated.
2. Click **Instance Configuration** and view the value of the DataNode parameter **dfs.datanode.data.dir**.
3. Check whether all DataNode data directories are consistent with the disk plan.
   - If yes, go to **Step 2.4**.
   - If no, go to **Step 2.7**.
4. Modify the DataNode parameter **dfs.datanode.data.dir** and delete the incorrect directories.
5. Choose **Components** > **HDFS** > **Instances** to restart the DataNode instance.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.7**.
7. Log in to the DataNode for which the alarm is generated.
   - If the alarm cause is "The DataNode data directory fails to be created", go to **Step 3.1**.
   - If the alarm cause is "The DataNode data directory uses the same disk as critical directories, such **/** or **/boot**", go to **Step 4.1**.
   - If the alarm cause is "Multiple directories in the DataNode data directory use the same disk", go to **Step 5.1**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the DataNode data directory fails to be created.

1. Run the following commands to switch the user:

   sudo su - root

   su - omm

2. Run the **ls** command to check whether the directories exist in the DataNode data directory.
   - If yes, go to **Step 7**.
   - If no, go to **Step 3.3**.
3. Run the **mkdir** *data directory* command to create a directory and check whether the directory is successfully created.
   - If yes, go to **Step 6.1**.
   - If no, go to **Step 3.4**.
4. Click **Alarms** to check whether alarm ALM-12017 Insufficient Disk Capacity exists.
   - If yes, go to **Step 3.5**.
   - If no, go to **Step 3.6**.
5. Adjust the disk capacity and check whether alarm ALM-12017 Insufficient Disk Capacity is cleared. For details, see **ALM-12017 Insufficient Disk Capacity (For MRS 2.x or Earlier)**.
   - If yes, go to **ALM-12017 Insufficient Disk Capacity (For MRS 2.x or Earlier)**.
   - If no, go to **Step 7**.
6. Check whether user **omm** has the **rwx** or **x** permission of all the upper-layer directories of the directory. (For example, for **/tmp/abc/**, user **omm** has the **x** permission for directory **tmp** and the **rwx** permission for directory **abc**.)
   - If yes, go to **Step 6.1**.
   - If no, go to **Step 3.7**.
7. Run the **chmod u+rwx** *path* or **chmod u+x** *path* command as the **root** user to add the **rwx** or **x** permission to the paths. Then, go to **Step 3.3**.

Step 4 Check whether the DataNode data directory uses the same disk as other critical directories in the system.

1. Run the **df** command to obtain the disk mounting information of each directory in the DataNode data directory.
2. Check whether the directories mounted to the disk are critical directories, such as **/** or **/boot**.
   - If yes, go to **Step 4.3**.
   - If no, go to **Step 6.1**.
3. Change the value of the DataNode parameter **dfs.datanode.data.dir** and delete the directories that use the same disk as critical directories.
4. Go to **Step 6.1**.

Step 5 Check whether multiple directories in the DataNode data directory use the same disk.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Run the **df** command to obtain the disk mounting information of each directory in the DataNode data directory. Record the mounted directory in the command output.
2. Modify the DataNode node parameter **dfs.datanode.data.dir** to reserve one of the directories mounted on the same disk directory.
3. Go to **Step 6.1**.

Step 6 Restart the DataNode and check whether the alarm is cleared.

1. Choose **Components** > **HDFS** > **Instances** to restart the DataNode instance.
2. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 7**.

Step 7 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.501 ALM-14012 HDFS Journalnode Data Is Not Synchronized (For MRS 2.*x* or Earlier)

Description

On the active NameNode, the system checks data synchronization on all JournalNodes in the cluster every 5 minutes. This alarm is generated when data on a JournalNode is not synchronized with that on other JournalNodes.

This alarm is cleared in 5 minutes after data on JournalNodes is synchronized.

Attribute

Alarm ID	Alarm Severity	Auto Clear
14012	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
IP	Specifies the service IP address of the JournalNode instance for which the alarm is generated.

Impact on the System

When a JournalNode is working incorrectly, data on the node is not synchronized with that on other JournalNodes. If data on more than half of JournalNodes is not synchronized, the NameNode cannot work correctly, making the HDFS service unavailable.

Possible Causes

- The JournalNode instance has not been started or has been stopped.
- The JournalNode instance is working incorrectly.
- The network of the JournalNode is unreachable.

Procedure

Step 1 Check whether the JournalNode instance has been started.

1. On the MRS cluster details page, click **Alarms**. In the alarm list, click the alarm.
2. In the **Alarm Details** area, check **Location** and obtain the IP address of the JournalNode for which the alarm is generated.
3. Choose **Components** > **HDFS** > **Instances**. In the instance list, click the JournalNode for which the alarm is generated and check whether **Operating Status** of the node is **Started**.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.4**.
4. Select the JournalNode instance and choose **More** > **Start Instance** to start it.
5. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 2 Check whether the JournalNode instance is working correctly.

1. Check whether **Health Status** of the JournalNode instance is **Good**.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 2.2**.
2. Select the JournalNode instance and choose **More** > **Restart Instance** to restart it.
3. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Check whether the network of the JournalNode is reachable.

1. On the MRS cluster details page, choose **Components** > **HDFS** > **Instances** to check the service IP address of the active NameNode.
2. Log in to the active NameNode.
3. Run the **ping** command to check whether a timeout occurs or the network between the active NameNode and the JournalNode is unreachable.

   ping *service IP address of the JournalNode*

   - If yes, go to **Step 3.4**.
   - If no, go to **Step 4**.
4. Contact O&M personnel to rectify the network fault. Wait 5 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.502 ALM-16000 Percentage of Sessions Connected to the HiveServer to the Maximum Number Allowed Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the percentage of sessions connected to the HiveServer to the maximum number allowed every 30 seconds. This indicator can be viewed on the Hive service monitoring page. This alarm is generated when the the percentage of sessions connected to the HiveServer to the maximum number allowed exceeds the specified threshold (90% by default).

This alarm can be automatically cleared when the percentage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16000	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

If a connection alarm is generated, too many sessions are connected to the HiveServer and new connections cannot be created.

Possible Causes

Too many clients are connected to the HiveServer.

Procedure

Step 1 Increase the maximum number of connections to Hive.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **Hive** > **Service Configuration** and switch **Basic** to **All**.
3. Increase the value of the **hive.server.session.control.maxconnections** configuration item. Suppose the value of the configuration item is A, the threshold is B, and sessions connected to the HiveServer is C. Adjust the value of the configuration item according to A x B > C. Sessions connected to the HiveServer can be viewed on the Hive monitoring page.
4. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.503 ALM-16001 Hive Warehouse Space Usage Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the Hive warehouse space usage every 30 seconds. The indicator **Percentage of HDFS Space Used by Hive to the Available Space** can be viewed on the Hive service monitoring page. This alarm is generated when the Hive warehouse space usage exceeds the specified threshold (85% by default).

This alarm is cleared when the Hive warehouse space usage is less than or equal to the threshold. You can reduce the warehouse space usage by expanding the warehouse capacity or releasing the used space.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16001	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

The system fails to write data, which causes data loss.

Possible Causes

- The upper limit of the HDFS capacity available for Hive is too small.
- The system disk space is insufficient.
- Some data nodes break down.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Expand the system configuration.

1. Analyze the cluster HDFS capacity usage and increase the upper limit of the HDFS capacity available for Hive.

   Go to the MRS cluster details page, choose **Components** > **Hive** > **Service Configuration**, set **Type** to **All**, search for **hive.metastore.warehouse.size.percent**, and increase the value of this parameter. Suppose that the value of the configuration item is A, total HDFS storage space is B, the threshold is C, and HDFS space used by Hive is D. Adjust the value of the configuration item according to A x B x C > D. The total HDFS storage space can be viewed on the HDFS monitoring page, and HDFS space used by Hive can be viewed on the Hive monitoring page.

2. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Expand the system.

1. Add nodes.
2. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3.1**.

Step 3 Check whether the data node is normal.

1. Go to the cluster details page and choose **Alarms**.
2. Check whether ALM-12006 Node Fault, ALM-12007 Process Fault, or ALM-14002 DataNode Disk Usage Exceeds the Threshold exists.
   - If yes, go to **Step 3.3**.
   - If no, go to **Step 4**.
3. Handle the alarm by following the steps provided in **ALM-12006 Node Fault (For MRS 2.x or Earlier)**, **ALM-12007 Process Fault (For MRS 2.x or Earlier)**, and **ALM-14002 DataNode Disk Usage Exceeds the Threshold (For MRS 2.x or Earlier)**.
4. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.504 ALM-16002 Hive SQL Execution Success Rate Is Lower Than the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the percentage of the HiveQL statements that are executed successfully every 30 seconds. Percentage of HiveQL statements that are executed successfully = Number of HiveQL statements that are executed successfully by Hive in a specified period/Total number of HiveQL statements that are executed by Hive. This indicator can be viewed on the Hive service monitoring page. This alarm is generated when the percentage of the HiveQL statements that are executed successfully exceeds the specified threshold (90% by default). The name of the host for which the alarm is generated can be obtained from the location information of the alarm. The host IP address is the IP address of the HiveServer node.

This alarm is cleared when the percentage of the HiveQL statements that are executed successfully in a test period is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16002	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm.

Impact on the System

The system configuration and performance cannot meet service processing requirements.

Possible Causes

- A syntax error occurs in HiveQL commands.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The HBase service is abnormal when a Hive on HBase task is being performed.
- Basic services that are depended on are abnormal, such as HDFS, Yarn, and ZooKeeper.

Procedure

Step 1 Check whether the HiveQL commands comply with syntax.

1. Use the Hive client to log in to the HiveServer node for which the alarm is generated. Query the HiveQL syntax standard provided by Apache, and check whether the HiveQL commands are correct.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.2**.

   NOTE

   To view the user who runs an incorrect statement, download HiveServerAudit logs of the HiveServer node for which this alarm is generated. Set **Start time** and **End time** to 10 minutes before and after the alarm generation time respectively. Open the log file and search for the **Result=FAIL** keyword to filter the log information about the incorrect statement, and then view the user who runs the incorrect statement according to **UserName** in the log information.
2. Enter correct HiveQL statements, and check whether the command can be properly executed.
   - If yes, go to **Step 4.5**.
   - If no, go to **Step 2.1**.

Step 2 Check whether the HBase service is abnormal.

1. Check whether a Hive on HBase task is performed.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 3.1**.
2. Check whether the HBase service is normal in the service list.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 2.3**.
3. Check the alarms displayed on the alarm page and clear them according to **Alarm Help**.
4. Enter correct HiveQL statements, and check whether the command can be properly executed.
   - If yes, go to **Step 4.5**.
   - If no, go to **Step 3.1**.

Step 3 Check whether the Spark service is abnormal.

1. Check whether the Spark service is normal in the service list.
   - If yes, go to **Step 4.1**.
   - If no, go to **Step 3.2**.
2. Check the alarms displayed on the alarm page and clear them according to **Alarm Help**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Enter correct HiveQL statements, and check whether the command can be properly executed.
   - If yes, go to **Step 4.5**.
   - If no, go to **Step 4.1**.

Step 4 Check whether HDFS, Yarn, and ZooKeeper are normal.

1. Go to the MRS cluster details page and click **Components**.
2. In the service list, check whether the services, such as HDFS, Yarn, and ZooKeeper are normal.
   - If yes, go to **Step 4.5**.
   - If no, go to **Step 4.3**.
3. Check the alarms displayed on the alarm page and clear them according to **Alarm Help**.
4. Enter correct HiveQL statements, and check whether the command can be properly executed.
   - If yes, go to **Step 4.5**.
   - If no, go to **Step 5**.
5. Wait one minute and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.505 ALM-16004 Hive Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Hive service status every 30 seconds. This alarm is generated when the Hive service is unavailable.

This alarm is cleared when the Hive service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16004	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The system cannot provide data loading, query, and extraction services.

Possible Causes

- Basic services, such as ZooKeeper, HDFS, Yarn, and DBService work incorrectly, or the Hive process is faulty.
  - ZooKeeper is abnormal.
  - HDFS is abnormal.
  - Yarn is abnormal.
  - DBService is abnormal.
  - The Hive service process is faulty. If the alarm is caused by a Hive process fault, the alarm report has a delay of about 5 minutes.
- The network communication between the Hive service and basic services is interrupted.

Procedure

Step 1 Check the HiveServer/MetaStore process status.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **Hive** > **Instances**. In the Hive instance list, check whether the status of all HiveServer and MetaStore instances is **Unknown**.
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 2**.
3. Above the Hive instance list, choose **More** > **Restart Instance** to restart the HiveServer/MetaStore process.
4. In the alarm list, check whether ALM-16004 Hive Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Check the ZooKeeper status.

1. Go to the MRS cluster details page and choose **Alarms**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. On MRS Manager, check whether the ALM-12007 Process Fault alarm is reported.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
3. In the **Alarm Details** area of ALM-12007 Process Fault, check whether **ServiceName** is **ZooKeeper**.
   - If yes, go to **Step 2.4**.
   - If no, go to **Step 3**.
4. Rectify the fault by following steps provided in **ALM-12007 Process Fault (For MRS 2.x or Earlier)**.
5. In the alarm list, check whether ALM-16004 Hive Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Check the HDFS status.

1. Go to the MRS cluster details page and choose **Alarms**.
2. In the alarm list, check whether the alarm ALM-14000 HDFS Service Unavailable exists.
   - If yes, go to **Step 3.3**.
   - If no, go to **Step 4**.
3. Rectify the fault by following the steps provided in **ALM-14000 HDFS Service Unavailable (For MRS 2.x or Earlier)**.
4. In the alarm list, check whether ALM-16004 Hive Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Check the Yarn status.

1. Go to the MRS cluster details page and choose **Alarms**.
2. In the alarm list on MRS Manager, check whether the alarm ALM-18000 Yarn Service Unavailable is generated.
   - If yes, go to **Step 4.3**.
   - If no, go to **Step 4**.
3. Rectify the fault by following the steps provided in **ALM-18000 Yarn Service Unavailable (For MRS 2.x or Earlier)**.
4. In the alarm list, check whether ALM-16004 Hive Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 5 Check the DBService status.

1. Go to the MRS cluster details page and choose **Alarms**.
2. In the alarm list on MRS Manager, check whether ALM-27001 DBService Unavailable is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 5.3**.
- If no, go to **Step 6**.

3. Rectify the fault by following the handling procedure in **ALM-27001 DBService Unavailable (For MRS 2.x or Earlier)**.
4. In the alarm list, check whether ALM-16004 Hive Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 6**.

Step 6 Check the network connection between Hive and ZooKeeper, HDFS, Yarn, and DBService.

1. Go to the MRS cluster details page and click **Components**.
2. Click **Hive**.
3. Click **Instances**.

   The HiveServer instance list is displayed.
4. Click **Host Name** in the row of **HiveServer**.

   The HiveServer host status page is displayed.
5. Record the IP address under **Summary**.
6. Use the IP address obtained in **Step 6.5** to log in to the host where HiveServer is located.
7. Run the **ping** command to check whether the network connection between the host that runs HiveServer and the hosts that run the ZooKeeper, HDFS, Yarn, and DBService services is normal. Methods of obtaining IP addresses of the hosts that run ZooKeeper, HDFS, Yarn, and DBService services as well as the HiveServer IP address are the same.
   - If yes, go to **Step 7**.
   - If no, go to **Step 6.8**.
8. Contact the O&M personnel to restore the network.
9. In the alarm list, check whether ALM-16004 Hive Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 7**.

Step 7 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

9.11.506 ALM-16005 Number of Failed Hive SQL Executions in the Last Period Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks whether the number of Hive SQL statements that fail to be executed has exceeded the threshold in the last 10-minute period. This alarm is generated when the number of failed Hive SQL executions in the last 10 minutes exceeds the threshold. In the next 10 minutes, if the number of failed Hive SQL statement executions is less than the threshold, the alarm is automatically cleared.

Attribute

Alarm ID	Alarm Severity	Auto Clear
16005	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

None

Possible Causes

The Hive SQL syntax is incorrect. As a result, the Hive SQL statements fail to be executed.

Procedure

Check the Hive SQL statements that fail to be executed, correct the syntax, and execute the SQL statements again.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.507 ALM-18000 Yarn Service Unavailable (For MRS 2.*x* or Earlier)

Description

The alarm module checks the Yarn service status every 30 seconds. This alarm is generated when the Yarn service is unavailable.

This alarm is cleared when the Yarn service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18000	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The cluster cannot provide the Yarn service. Users cannot run new applications. Submitted applications cannot be run.

Possible Causes

- ZooKeeper is abnormal.
- HDFS is abnormal.
- There is no active ResourceManager node in the Yarn cluster.
- All NodeManager nodes in the Yarn cluster are abnormal.

Procedure

Step 1 Check the ZooKeeper status.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, check whether the alarm ALM-13000 ZooKeeper Service Unavailable exists.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 1.3**.
- If no, go to **Step 2.2**.

3. Rectify the fault by following the handling procedure in **ALM-13000 ZooKeeper Service Unavailable (For MRS 2.x or Earlier)**. Then, check whether this alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.2**.

Step 2 Check the HDFS status.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, check whether an HDFS alarm is generated.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3.2**.
3. Click **Alarms**, and handle HDFS alarms according to **Alarm Help**. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3.2**.

Step 3 Check the ResorceManager status in the Yarn cluster.

1. Go to the MRS cluster details page and click **Components**.
2. Click **Yarn**.
3. In **Yarn Summary**, check whether there is an active ResourceManager node in the Yarn cluster.
   - If yes, go to **Step 4.2**.
   - If no, go to **Step 5**.

Step 4 Check the NodeManager node status in the Yarn cluster.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **Yarn** > **Instances**.
3. Check **Health Status** of NodeManager, and check whether there are unhealthy nodes.
   - If yes, go to **Step 4.4**.
   - If no, go to **Step 5**.
4. Rectify the fault by following the procedure provided in **ALM-18002 NodeManager Heartbeat Lost (For MRS 2.x or Earlier)** or **ALM-18003 NodeManager Unhealthy (For MRS 2.x or Earlier)**. Then, check whether this alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.508 ALM-18002 NodeManager Heartbeat Lost (For MRS 2.*x* or Earlier)

Description

The system checks the number of lost NodeManager nodes every 30 seconds, and compares the number of lost nodes with the threshold. The **Lost Nodes** indicator has a default threshold. This alarm is generated when the value of the **Lost Nodes** indicator exceeds the threshold.

This alarm is cleared when the value of **Lost Nodes** is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18002	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

- The lost NodeManager node cannot provide the Yarn service.
- The number of containers decreases, so the cluster performance deteriorates.

Possible Causes

- NodeManager is forcibly deleted without decommission.
- All NodeManager instances are stopped or the NodeManager process is faulty.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The host where the NodeManager node resides is faulty.
- The network between the NodeManager and ResourceManager is disconnected or busy.

Procedure

Step 1 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.509 ALM-18003 NodeManager Unhealthy (For MRS 2.*x* or Earlier)

Description

The system checks the number of abnormal NodeManager nodes every 30 seconds, and compares the number of abnormal nodes with the threshold. The **Unhealthy Nodes** indicator has a default threshold. This alarm is generated when the value of the **Unhealthy Nodes** indicator exceeds the threshold.

This alarm is cleared when the value of **Unhealthy Nodes** is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18003	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

- The faulty NodeManager node cannot provide the Yarn service.
- The number of containers decreases, so the cluster performance deteriorates.

Possible Causes

- The disk space of the host where the NodeManager node resides is insufficient.
- User **omm** does not have the permission to access a local directory on the NodeManager node.

Procedure

Step 1 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.510 ALM-18004 NodeManager Disk Usability Ratio Is Lower Than the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the available disk space of each NodeManager node every 30 seconds and compares the disk availability rate with the threshold. A default threshold range is provided for the **NodeManager Disk Usability Ratio**. This alarm is generated when the system detects that the actual **NodeManager Disk Usability Ratio** is lower than the threshold.

This alarm is automatically cleared when the value of **NodeManager Disk Usability Ratio** is greater than the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18004	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Specifies the threshold for triggering the alarm.

Impact on the System

- The NodeManager node whose disk availability rate is lower than the threshold may fail to provide the Yarn service.
- The number of containers decreases, so the cluster performance may deteriorate.

Possible Causes

- The disk space of the host where the NodeManager node resides is insufficient.
- User **omm** does not have the permission to access a local directory on the NodeManager node.

Procedure

Step 1 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.511 ALM-18006 MapReduce Job Execution Timeout (For MRS 2.*x* or Earlier)

Description

The alarm module checks the MapReduce job execution every 30 seconds. This alarm is generated when the execution of a submitted MapReduce job times out.

This alarm must be manually cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18006	Major	No

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Execution of the submitted MapReduce job times out, so no execution result can be obtained. Execute the job again after rectifying the fault.

Possible Causes

It takes a long time to execute a MapReduce job. However, the specified time is less than the required execution time.

Procedure

Step 1 Check whether time is improperly set.

Set **-Dapplication.timeout.interval** to a larger value, or do not set the parameter. Check whether the MapReduce job can be executed.

- If yes, go to **Step 2.5**.
- If no, go to **Step 2.2**.

Step 2 Check the Yarn status.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list on MRS Manager, check whether the alarm ALM-18000 Yarn Service Unavailable is generated.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Rectify the fault by following the handling procedure in **ALM-18000 Yarn Service Unavailable (For MRS 2.x or Earlier)**.
4. Run the MapReduce job command again to check whether the MapReduce job can be executed.
   - If yes, go to **Step 2.5**.
   - If no, go to **Step 4**.
5. In the alarm list, click  in the **Operation** column of the alarm to manually clear the alarm. No further action is required.

Step 3 Adjust the timeout threshold.

On MRS Manager, choose **System** > **Threshold Configuration** > **Services** > **Yarn** > **Timed out Applications**, and increase the maximum number of timeout tasks allowed by the current threshold rule. Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.512 ALM-18008 Heap Memory Usage of Yarn ResourceManager Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the heap memory usage of Yarn ResourceManager every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the heap memory usage of Yarn ResourceManager exceeds the threshold (80% of the maximum memory by default).

To change the threshold, choose **System** > **Threshold Configuration** > **Service** > **Yarn**. The alarm is cleared when the heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18008	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

When the heap memory usage of Yarn ResourceManager is overhigh, the performance of Yarn task submission and operation is affected. What is more, a memory overflow occurs so that the Yarn service is unavailable.

Possible Causes

The heap memory of the Yarn ResourceManager instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Step 1 Check the heap memory usage.

1. Go to the MRS cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **18008** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Yarn** > **Instances** > **ResourceManager** (IP address of the instance for which the alarm is generated) > **Customize** > **Percentage of Used Heap Memory of the ResourceManager**. Check the heap memory usage.
4. Check whether the heap memory usage of ResourceManager has reached the threshold (80% of the maximum memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Yarn** > **Service Configuration**. Set **Type** to **All** and choose **ResourceManager** > **System**. Change the values of **-Xmx** and **-Xms** in the **GC_OPTS** parameter based on the site requirements to ensure that the value of **-Xms** is less than that of **-Xmx**. Click **Save Configuration** and select **Restart Role Instance**. Click **OK**.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.513 ALM-18009 Heap Memory Usage of MapReduce JobHistoryServer Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the heap memory usage of MapReduce JobHistoryServer every 30 seconds and compares the actual usage with the threshold. The alarm is generated when the heap memory usage of MapReduce JobHistoryServer exceeds the threshold (80% of the maximum memory by default).

To change the threshold, choose **System** > **Threshold Configuration** > **Service** > **MapReduce**. The alarm is cleared when the heap memory usage is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18009	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

When the heap memory usage of MapReduce JobHistoryServer is overhigh, the performance of MapReduce log archiving is affected. What is more, a memory overflow occurs so that the Yarn service is unavailable.

Possible Causes

The heap memory of the MapReduce JobHistoryServer instance on the node is overused or the heap memory is inappropriately allocated. As a result, the usage exceeds the threshold.

Procedure

Step 1 Check the heap memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **18009** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **MapReduce** > **Instance** > **JobHistoryServer** (IP address of the instance for which the alarm is generated) > **Customize** > **JobHistoryServer Heap Memory Usage Statistics**. Check the heap memory usage.
4. Check whether the heap memory usage of JobHistoryServer has reached the threshold (80% of the maximum heap memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **MapReduce** > **Service Configuration**. Set **Type** to **All** and choose **JobHistoryServer** > **System**. Increase the value of **-Xmx** in the **GC_OPTS** parameter as required, click **Save Configuration**, and select **Restart the affected services or instances.** Click **OK**.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.514 ALM-18010 Number of Pending Yarn Tasks Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the number of pending Yarn tasks every 30 seconds and compares the number of tasks with the threshold. This alarm is generated when the number of pending tasks exceeds the threshold.

You can change the threshold by choosing **System** > **Configure Alarm Threshold** > **Service** > **Yarn** > **Queue Root Pending Applications** > **Queue Root Pending Applications** on MRS Manager.

This alarm is cleared when the number of pending tasks is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18010	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Tasks may be stacked and cannot be processed in a timely manner.

Possible Causes

The computing capability of the cluster is lower than the task submission rate. As a result, the task cannot be processed in a timely manner after being submitted.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Check the usage of memory and vCores on the Yarn page.

Check whether the values of **Memory Used|Memory Total** and **VCores Used|VCores Total** on the native Yarn page reach or approach the maximum values.

- If yes, go to **Step 2**.
- If no, go to **Step 5**.

Step 2 Check the number of submitted tasks.

Check whether the running tasks are submitted at a normal frequency.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Scale out the cluster.

The scale-out is based on the site requirements. For details, see .

Step 4 After the scale-out is completed, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.515 ALM-18011 Memory of Pending Yarn Tasks Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the memory of pending Yarn tasks every 30 seconds and compares the memory with the threshold. This alarm is generated when the memory of pending tasks exceeds the threshold.

You can change the threshold by choosing **System** > **Configure Alarm Threshold** > **Service** > **Yarn** > **Queue Root Pending Memory** > **Queue Root Pending Memory** on MRS Manager.

This alarm is cleared when the memory of pending tasks is less than or equal to the threshold.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
18011	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Specifies the threshold for triggering the alarm.

Impact on the System

Tasks may be stacked and cannot be processed in a timely manner.

Possible Causes

The computing capability of the cluster is lower than the task submission rate. As a result, the task cannot be processed in a timely manner after being submitted.

Procedure

Step 1 Check the usage of memory and vCores on the Yarn page.

Check whether the values of **Memory Used|Memory Total** and **VCores Used|VCores Total** on the native Yarn page reach or approach the maximum values.

- If yes, go to **Step 2**.
- If no, go to **Step 5**.

Step 2 Check the number of submitted tasks.

Check whether the running tasks are submitted at a normal frequency.

- If yes, go to **Step 3**.
- If no, go to **Step 5**.

Step 3 Scale out the cluster.

The scale-out is based on the site requirements. For details, see .

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 After the scale-out is completed, check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Step 5 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.516 ALM-18012 Number of Terminated Yarn Tasks in the Last Period Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the number of terminated Yarn tasks every 10 minutes. This alarm is generated when the number of terminated Yarn tasks in the last 10 minutes is greater than the threshold. This alarm is automatically cleared when the number of terminated Yarn tasks is less than the threshold in the next 10 minutes.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18012	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

None

Possible Causes

A user manually stops a running Yarn task.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Check the task termination operator in the Yarn logs and audit logs, and determine the cause of the task termination.

Reference

None

9.11.517 ALM-18013 Number of Failed Yarn Tasks in the Last Period Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the number of failed Yarn tasks every 10 minutes. This alarm is generated when the number of failed Yarn tasks in the last 10 minutes is greater than the threshold. This alarm is automatically cleared when the number of failed Yarn tasks is less than the threshold in the next 10 minutes.

Attribute

Alarm ID	Alarm Severity	Auto Clear
18013	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

None

Possible Causes

The submitted Yarn job program is incorrect. For example, the parameter for Spark to submit a job is incorrect.

Procedure

Check the log of the failed job, locate the failure cause, modify the job, and submit the job again.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.518 ALM-19000 HBase Service Unavailable (For MRS 2.*x* or Earlier)

Description

The alarm module checks the HBase service status every 30 seconds. This alarm is generated when the HBase service is unavailable.

This alarm is cleared when the HBase service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
19000	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Operations cannot be performed, such as reading or writing data and creating tables.

Possible Causes

- ZooKeeper is abnormal.
- HDFS is abnormal.
- HBase is abnormal.
- The network is abnormal.

Procedure

Step 1 Check the ZooKeeper status.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Go to the MRS cluster details page and click **Components**.
2. In the service list, check whether the health status of ZooKeeper is **Good**.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.3**.
3. In the alarm list, check whether the alarm ALM-13000 ZooKeeper Service Unavailable exists.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2.1**.
4. Rectify the fault by following the steps provided in **ALM-13000 ZooKeeper Service Unavailable (For MRS 2.x or Earlier)**.
5. Wait several minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the HDFS status.

1. On MRS Manager, check whether the ALM-14000 HDFS Service Unavailable alarm is reported.
   - If yes, go to **Step 2.2**.
   - If no, go to **Step 3**.
2. Rectify the fault by following the steps provided in **ALM-14000 HDFS Service Unavailable (For MRS 2.x or Earlier)**.
3. Wait several minutes and check whether the alarm is cleared.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Reference

None

9.11.519 ALM-19006 HBase Replication Sync Failed (For MRS 2.*x* or Earlier)

Description

This alarm is generated when disaster recovery (DR) data fails to be synchronized to a standby cluster.

This alarm is cleared when DR data synchronization succeeds.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Attribute

Alarm ID	Alarm Severity	Auto Clear
19006	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

HBase data in a cluster fails to be synchronized to the standby cluster, causing data inconsistency between active and standby clusters.

Possible Causes

- The HBase service on the standby cluster is abnormal.
- The network is abnormal.

Procedure

Step 1 Observe whether the system automatically clears the alarm.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, click the alarm to obtain alarm generation time from **Generated Time** in **Alarm Details**. Check whether the alarm has existed for over 5 minutes.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.3**.
3. Wait 5 minutes and check whether the alarm is automatically cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the HBase service status of the standby cluster.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, click the alarm and obtain **HostName** from **Location** in **Alarm Details**.
3. Log in to the node where the HBase client of the active cluster is located. Run the following commands to switch the user:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

sudo su - root

su - omm

4. Run the **status 'replication', 'source'** command to check the synchronization status of the faulty node.

   The synchronization status of a node is as follows.

   ```
   10-10-10-153:
    SOURCE: PeerID=abc, SizeOfLogQueue=0, ShippedBatches=2, ShippedOps=2, ShippedBytes=320,
   LogReadInBytes=1636, LogEditsRead=5, LogEditsFiltered=3, SizeOfLogToReplicate=0,
   TimeForLogToReplicate=0, ShippedHFiles=0, SizeOfHFileRefsQueue=0, AgeOfLastShippedOp=0,
   TimeStampsOfLastShippedOp=Mon Jul 18 09:53:28 CST 2016, Replication Lag=0,
   FailedReplicationAttempts=0
    SOURCE: PeerID=abc1, SizeOfLogQueue=0, ShippedBatches=1, ShippedOps=1, ShippedBytes=160,
   LogReadInBytes=1636, LogEditsRead=5, LogEditsFiltered=3, SizeOfLogToReplicate=0,
   TimeForLogToReplicate=0, ShippedHFiles=0, SizeOfHFileRefsQueue=0, AgeOfLastShippedOp=16788,
   TimeStampsOfLastShippedOp=Sat Jul 16 13:19:00 CST 2016, Replication Lag=16788,
   FailedReplicationAttempts=5
   ```

5. Obtain **PeerID** corresponding to a record whose **FailedReplicationAttempts** value is greater than 0.

   In the preceding step, data on the faulty node **10-10-10-153** fails to be synchronized to a standby cluster whose **PeerID** is **abc1**.

6. Run the **list_peers** command to find the cluster and the HBase instance corresponding to **PeerID**.

   ```
   PEER_ID CLUSTER_KEY STATE TABLE_CFS
    abc1 10.10.10.110,10.10.10.119,10.10.10.133:24002:/hbase2 ENABLED
    abc 10.10.10.110,10.10.10.119,10.10.10.133:24002:/hbase ENABLED
   ```

   In the preceding information, **/hbase2** indicates that data is synchronized to the HBase2 instance of the standby cluster.

7. In the service list of the standby cluster, check whether the health status of the HBase instance obtained in **Step 2.6** is **Good**.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 2.8**.
8. In the alarm list, check whether the alarm ALM-19000 HBase Service Unavailable exists.
   - If yes, go to **Step 2.9**.
   - If no, go to **Step 3.1**.
9. Rectify the fault by following the steps provided in **ALM-19000 HBase Service Unavailable (For MRS 2.x or Earlier)**.
10. Wait several minutes and check whether the alarm is cleared.
    - If yes, no further action is required.
    - If no, go to **Step 3.1**.

Step 3 Check the network connection between RegionServers on active and standby clusters.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, click the alarm and obtain **HostName** from **Location** in **Alarm Details**.
3. Log in to the faulty RegionServer node.
4. Run the **ping** command to check whether the network connection between the faulty RegionServer node and the host where RegionServer of the standby cluster resides is normal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 4**.
- If no, go to **Step 3.5**.

5. Contact the O&M personnel to restore the network.
6. After the network recovers, check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Reference

None

9.11.520 ALM-19007 HBase Merge Queue Exceeds the Threshold (for 2.x and Earlier Versions)

Description

The system checks the HBase compaction queue size every 30 seconds. This alarm is generated when the compaction queue size exceeds the alarm threshold (**100** by default) for three consecutive times. This alarm is cleared when the compaction queue size is less than the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
19007	Minor	Yes

Parameters

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
Host Name	Specifies the host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

The cluster performance may deteriorate, affecting data read and write.

Possible Causes

- The number of HBase RegionServers is too small.
- There are too many regions on a RegionServer of HBase.
- The HBase RegionServer heap size is small.
- Resources are insufficient.
- Related parameters are not configured properly.

Procedure

Step 1 Check whether related HBase parameters are properly configured.

1. Log in to the MRS cluster details page, choose **Components** > **HBase** > **Service Configuration**, switch **Basic Configuration** to **All Configurations**, and search for **hbase.hstore.compaction.min** and **hbase.hstore.compaction.max**, and increase the values of **hbase.regionserver.thread.compaction.small** and **hbase.regionserver.thread.compaction.throttle**.

   NOTE

   If you did not synchronize IAM users, perform synchronization first. (In the **Dashboard** tab, click **Synchronize** next to **IAM User Sync**.)

2. Save the configuration, and restart the HBase service during off-peak hours or perform a rolling restart to make the configuration take effect.
3. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.521 ALM-20002 Hue Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Hue service status every 60 seconds. This alarm is generated if the Hue service is unavailable.

This alarm is cleared when the Hue service is normal.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
20002	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The system cannot provide data loading, query, and extraction services.

Possible Causes

- The KrbServer service on which Hue depends is abnormal.
- The DBService service on which Hue depends is abnormal.
- The network connection to DBService is abnormal.

Procedure

Check whether the KrbServer service is normal.

Step 1 Go to the MRS cluster details page and click **Components**.

Step 2 In the service list, check whether **Health Status** of **KrbServer** is **Good**.

- If yes, go to **Step 5**.
- If no, go to **Step 3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Click **Restart** in the **Operation** column of the KrbServer service to restart the service.

Step 4 Wait for several minutes. Check whether ALM-20002 Hue Service Unavailable is cleared.

- If yes, no further action is required.
- If no, go to **Step 5**.

Check whether DBService is normal.

Step 5 Go to the MRS cluster details page and click **Components**.

Step 6 In the service list, check whether **Health Status** of **DBService** is **Good**.

- If yes, go to **Step 9**.
- If no, go to **Step 7**.

Step 7 Click **Restart** in the **Operation** column of the DBService service to restart the service.

NOTE

To restart the service, you need to enter the password of the MRS Manager administrator and select **Start or restart related services** .

Step 8 Wait for several minutes. Check whether ALM-20002 Hue Service Unavailable is cleared.

- If yes, no further action is required.
- If no, go to **Step 9**.

Check whether the network connected to DBService is normal.

Step 9 Choose **Components** > **Hue** > **Instance** and record the IP address of the active Hue node.

Step 10 Use PuTTY to log in to the active Hue.

Step 11 Run the **ping** command to check whether the network connection between the host where the active Hue is located and the host where DBService is located is normal. (The method of obtaining the DBService service IP address is the same as that of obtaining the active Hue IP address.)

- If yes, go to **Step 17**.
- If no, go to **Step 12**.

Step 12 Contact the network administrator to repair the network.

Step 13 Wait for several minutes. Check whether ALM-20002 Hue Service Unavailable is cleared.

- If yes, no further action is required.
- If no, go to **Step 17**.

Collect fault information.

Step 14 On MRS Manager, choose **System** > **Export Log**.

Step 15 Select the following nodes from the **Services** drop-down list and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Hue
- Controller

Step 16 Set **Start Time** and **End Time** for log collection to 10 minutes before and after the alarm is generated, select an export type, and click **OK** to collect the corresponding fault log information.

Restart Hue.

Step 17 Choose **Components** > **Hue**.

Step 18 Choose **More** > **Restart Service** and click **OK**.

Step 19 Check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 20**.

Step 20 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.522 ALM-23001 Loader Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Loader service availability every 60 seconds. This alarm is generated if the Loader service is unavailable and is cleared after the Loader service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
23001	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Data loading, import, and conversion are unavailable.

Possible Causes

- The services that Loader depends on are abnormal.
  - ZooKeeper is abnormal.
  - HDFS is abnormal.
  - DBService is abnormal.
  - Yarn is abnormal.
  - MapReduce is abnormal.
- The network is faulty. Loader cannot communicate with its dependent services.
- Loader is running improperly.

Procedure

Step 1 Check the ZooKeeper status.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **ZooKeeper** and check whether the health status of ZooKeeper is normal.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 1.3**.
3. Choose **More > Restart Service** to restart ZooKeeper. After ZooKeeper starts, check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 1.4**.
4. On MRS Manager, check whether the ALM-12007 Process Fault alarm is reported.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2.1**.
5. In **Alarm Details** of the "ALM-12007 Process Fault" alarm, check whether **ServiceName** is **ZooKeeper**.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 2.1**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

6. Rectify the fault by following steps provided in **ALM-12007 Process Fault (For MRS 2.x or Earlier)**.
7. Check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check the HDFS status.

1. Go to the MRS cluster details page and choose **Alarms**.
2. On MRS Manager, check whether the "ALM-14000 HDFS Service Unavailable alarm" is reported.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3.1**.
3. Rectify the fault by following the steps provided in **ALM-14000 HDFS Service Unavailable (For MRS 2.x or Earlier)**.
4. Check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3.1**.

Step 3 Check the DBService status.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **DBService** to check whether the health status of DBService is normal.
   - If yes, go to **Step 4.1**.
   - If no, go to **Step 3.3**.
3. Choose **More > Restart Service** to restart DBService. After DBService starts, check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4.1**.

Step 4 Check the MapReduce status.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **MapReduce** and check whether the health status of MapReduce is normal.
   - If yes, go to **Step 5.1**.
   - If no, go to **Step 4.3**.
3. Choose **More > Restart Service** to restart MapReduce. After MapReduce starts, check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5.1**.

Step 5 Check the Yarn status.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **Yarn** and check whether the health status of Yarn is normal.
   - If yes, go to **Step 5.4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 5.3**.

3. Choose **More > Restart Service** to restart Yarn. After Yarn starts, check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 5.4**.
4. On MRS Manager, check whether the "ALM-18000 Yarn Service Unavailable" alarm is reported.
   - If yes, go to **Step 5.5**.
   - If no, go to **Step 6.1**.
5. Rectify the fault by following the steps provided in **ALM-18000 Yarn Service Unavailable (For MRS 2.x or Earlier)**.
6. Check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 6.1**.

Step 6 Check the network connections between Loader and its dependent components.

1. Go to the MRS cluster details page and click **Components**.
2. Click **Loader**.
3. Click **Instance**. The Sqoop instance list is displayed.
4. Record the management IP addresses of all Sqoop instances.
5. Log in to the hosts using the IP addresses obtained in **Step 6.4**. Run the following commands to switch the user:

   sudo su - root

   su - omm
6. Run the **ping** command to check whether the network connection between the hosts where the Sqoop instances reside and the dependent components is normal. (The dependent components include ZooKeeper, DBService, HDFS, MapReduce, and Yarn. The method to obtain the IP addresses of the dependent components is the same as that used to obtain the IP addresses of the Sqoop instances.)
   - If yes, go to **Step 7**.
   - If no, go to **Step 6.7**.
7. Contact the network administrator to repair the network.
8. Check whether the "ALM-23001 Loader Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 7**.

Step 7 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.523 ALM-24000 Flume Service Unavailable (For MRS 2.*x* or Earlier)

Description

The alarm module checks the Flume service status every 180 seconds. This alarm is generated if the Flume service is abnormal.

This alarm is cleared after the Flume service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
24000	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Flume cannot work and data transmission is interrupted.

Possible Causes

- HDFS is unavailable.
- LdapServer is unavailable.

Procedure

Step 1 Check the HDFS status.

1. Go to the MRS cluster details page and choose **Alarms**.
2. Check whether the ALM-14000 HDFS Service Unavailable alarm is generated.
   - If yes, clear the alarm according to the handling suggestions of **ALM-14000 HDFS Service Unavailable (For MRS 2.x or Earlier)**.
   - If no, go to **Step 2**.

Step 2 Check the LdapServer status.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Check whether the ALM-25000 LdapServer Service Unavailable alarm is generated.

- If yes, clear the alarm according to the handling suggestions of **ALM-25000 LdapServer Service Unavailable (For MRS 2.x or Earlier)**.
- If no, go to **Step 3.2**.

Step 3 Check whether the HDFS and LdapServer services are stopped.

1. Go to the MRS cluster details page and click **Components**.
2. In the service list on MRS Manager, check whether the HDFS and LdapServer services are stopped.
   - If yes, start the HDFS and LdapServer services and go to **Step 3.3**.
   - If no, go to **Step 4**.
3. Check whether the "ALM-24000 Flume Service Unavailable" alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.524 ALM-24001 Flume Agent Is Abnormal (For MRS 2.*x* or Earlier)

Description

This alarm is generated if the Flume agent monitoring module detects that the Flume agent process is abnormal.

This alarm is cleared after the Flume agent process recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
24001	Minor	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Functions of the alarmed Flume agent instance are abnormal. Data transmission tasks of the instance are suspended. In real-time data transmission, data will be lost.

Possible Causes

- The **JAVA_HOME** directory does not exist or the Java permission is incorrect.
- The permission of the Flume agent directory is incorrect.

Procedure

Step 1 Check the Flume agent's configuration file.

1. Log in to the host where the faulty node resides. Run the following command to switch to user **root**:

   sudo su - root

2. Run the **cd** *Flume installation directory*/**fusioninsight-flume-1.6.0/conf/** command to go to Flume's configuration directory.
3. Run the **cat ENV_VARS** command. Check whether the **JAVA_HOME** directory exists and whether the Flume agent user has execute permission of Java.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.4**.
4. Specify the correct JAVA_HOME directory and grant the Flume agent user with the execute permission of Java. Then go to **Step 2.4**.

Step 2 Check the permission of the Flume agent directory.

1. Log in to the host where the faulty node resides. Run the following command to switch to user **root**:

   sudo su - root

2. Run the following command to access the installation directory of the Flume agent:

   cd *Flume agent installation directory*

3. Run the **ls -al * -R** command. Check whether the owner of all files is the Flume agent user.
   - If yes, go to **Step 3**.
   - If no, run the **chown** command and change the owner of the files to the Flume agent user. Then go to **Step 2.4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.525 ALM-24003 Flume Client Connection Interrupted (For MRS 2.*x* or Earlier)

Description

The alarm module monitors the port connection status on the Flume server. This alarm is generated if the Flume server fails to receive a connection message from the Flume client in 3 consecutive minutes.

This alarm is cleared after the Flume server receives a connection message from the Flume client.

Attribute

Alarm ID	Alarm Severity	Auto Clear
24003	Major	Yes

Parameters

Parameter	Description
ClientIP	Specifies the IP address of the Flume client.
ServerIP	Specifies the IP address of the Flume server.
ServerPort	Specifies the port on the Flume server.

Impact on the System

The communication between the Flume client and server fails. The Flume client cannot send data to the Flume server.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The network between the Flume client and server is faulty.
- The Flume client's process is abnormal.
- The Flume client is incorrectly configured.

Procedure

Step 1 Check the network between the Flume client and server.

1. Log in to the host where the alarmed Flume client resides. Run the following command to switch to user **root**:

   sudo su - root

2. Run the **ping** *Flume server IP address* command to check whether the network between the Flume client and server is normal.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 4**.

Step 2 Check whether the Flume client's process is normal.

1. Log in to the host where the alarmed Flume client resides. Run the following command to switch to user **root**:

   sudo su - root

2. Run the **ps -ef|grep flume |grep client** command to check whether the Flume client process exists.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 4**.

Step 3 Check the Flume client configuration.

1. Log in to the host where the alarmed Flume client resides. Run the following command to switch to user **root**:

   sudo su - root

2. Run the **cd** *Flume installation directory*/**fusioninsight-flume-1.6.0/conf/** command to go to Flume's configuration directory.
3. Run the **cat properties.properties** command to query the current configuration file of the Flume client.
4. Check whether the **properties.properties** file is correctly configured according to the configuration description of the Flume agent.
   - If yes, go to **Step 3.5**.
   - If no, go to **Step 4**.
5. Modify the **properties.properties** configuration file.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.526 ALM-24004 Flume Fails to Read Data (For MRS 2.*x* or Earlier)

Description

The alarm module monitors the Flume source status. This alarm is generated if the duration that Flume source fails to read data exceeds the threshold.

Users can modify the threshold as required.

This alarm is cleared if the source reads data successfully.

Attribute

Alarm ID	Alarm Severity	Auto Clear
24004	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
ComponentType	Specifies the component type for which the alarm is generated.
ComponentName	Specifies the component name for which the alarm is generated.

Impact on the System

Data collection is stopped.

Possible Causes

- The Flume source is faulty.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The network is faulty.

Procedure

Step 1 Check whether the Flume source is normal.

1. Check whether the Flume source is the spoolDir type.
   - If yes, go to **Step 1.2**.
   - If no, go to **Step 1.3**.
2. Query the **spoolDir** directory and check whether all files have been sent.
   - If yes, no further action is required.
   - If no, go to **Step 1.5**.
3. Check whether the Flume source is the Kafka type.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 1.5**.
4. Log in to the Kafka client and run the following commands to check whether all topic data configured for the Kafka source has been consumed.

   cd /opt/client/Kafka/kafka/bin

   ./kafka-consumer-groups.sh --bootstrap-server *Kafka cluster IP address*:**21007 --new-consumer --describe --group example-group1 --command-config**

   ../config/consumer.properties
   - If yes, no further action is required.
   - If no, go to **Step 1.5**.
5. Go to the cluster details page and click **Components**.
6. Choose **Flume** > **Instances**.
7. Click the Flume instance of the faulty node and check whether the value of the **Source Speed Metrics** is 0.
   - If yes, go to **Step 2.1**.
   - If no, no further action is required.

Step 2 Check the status of the network between the Flume source and faulty node.

1. Check whether the Flume source is the avro type.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
2. Log in to the host where the faulty node resides. Run the following command to switch to user **root**:

   sudo su - root
3. Run the **ping** *Flume source IP address* command to check whether the Flume source can be pinged.
   - If yes, go to **Step 3**.
   - If no, go to **Step 2.4**.
4. Contact the network administrator to repair the network.
5. Wait for a while and check whether the alarm is cleared.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, no further action is required.
- If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.527 ALM-24005 Data Transmission by Flume Is Abnormal (For MRS 2.*x* or Earlier)

Description

The alarm module monitors the capacity of Flume channels. This alarm is generated if the duration that a channel is full or the number of times that a source fails to send data to the channel exceeds the threshold.

Users can set the threshold as required by modifying the **channelfullcount** parameter.

This alarm is cleared after the Flume channel space is released.

Attribute

Alarm ID	Alarm Severity	Auto Clear
24005	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
ComponentType	Specifies the component type for which the alarm is generated.
ComponentName	Specifies the component name for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the usage of the Flume channel continues to grow, the data transmission time increases. When the usage reaches 100%, the Flume agent process is suspended.

Possible Causes

- The Flume sink is faulty.
- The network is faulty.

Procedure

Step 1 Check whether the Flume sink is normal.

1. Check whether the Flume sink is the HDFS type.
   - If yes, go to **Step 1.2**.
   - If no, go to **Step 1.3**.
2. On MRS Manager, check whether the ALM-14000 HDFS Service Unavailable alarm is reported and whether the HDFS service is stopped.
   - If the alarm is reported, clear it according to the handling suggestions of **ALM-14000 HDFS Service Unavailable (For MRS 2.x or Earlier)**; if the HDFS service is stopped, start it. Then go to **Step 1.7**.
   - If no, go to **Step 1.7**.
3. Check whether the Flume sink is the HBase type.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 1.7**.
4. On MRS Manager, check whether the ALM-19000 HBase Service Unavailable alarm is reported and whether the HBase service is stopped.
   - If the alarm is reported, clear it according to the handling suggestions of **ALM-19000 HBase Service Unavailable (For MRS 2.x or Earlier)**; if the HBase service is stopped, start it. Then go to **Step 1.7**.
   - If no, go to **Step 1.7**.
5. Check whether the Flume sink is the Kafka type.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 1.7**.
6. On MRS Manager, check whether the ALM-38000 Kafka Service Unavailable alarm is reported and whether the Kafka service is stopped.
   - If the alarm is reported, clear it according to the handling suggestions of **ALM-38000 Kafka Service Unavailable (For MRS 2.x or Earlier)**; if the Kafka service is stopped, start it. Then go to **Step 1.7**.
   - If no, go to **Step 1.7**.
7. Go to the MRS cluster details page and click **Components**.
8. Choose **Flume** > **Instances**.
9. Click the Flume instance of the faulty node and check whether the value of the **Sink Speed Metrics** is 0.
   - If yes, go to **Step 2.1**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, no further action is required.

Step 2 Check the status of the network between the Flume sink and faulty node.

1. Check whether the Flume sink is the Avro type.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
2. Log in to the host where the faulty node resides. Run the following command to switch to user **root**:

   sudo su - root
3. Run the **ping** *Flume sink IP address* command to check whether the Flume sink can be pinged.
   - If yes, go to **Step 3**.
   - If no, go to **Step 2.4**.
4. Contact the network administrator to repair the network.
5. Wait for a while and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.528 ALM-25000 LdapServer Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the LdapServer service status every 30 seconds. This alarm is generated when the active and standby LdapServer services are abnormal.

This alarm is cleared when either of the LdapServer services restores.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25000	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

When this alarm is generated, no operation can be performed for the KrbServer users and LdapServer users in the cluster. For example, users, user groups, or roles cannot be added, deleted, or modified, and user passwords cannot be changed on MRS Manager. The authentication for existing users in the cluster is not affected.

Possible Causes

- The node where the LdapServer service locates is faulty.
- The LdapServer process is abnormal.

Procedure

Step 1 Check whether the nodes where the two SlapdServer instances of the LdapServer service locate are faulty.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **LdapServer** > **Instances**. Go to the LdapServer instance page to obtain the host name of the node where the two SlapdServer instances reside.
3. On the **Alarms** page of MRS Manager, check whether the alarm ALM-12006 Node Fault is generated.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2.1**.
4. Check whether the host name in the alarm information is the same as the actual host name in **Step 1.2**.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2.1**.
5. Rectify the fault by following steps provided in ALM-12006 Node Fault.
6. In the alarm list, check whether the alarm ALM-25000 LdapServer Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 2 Check whether the LdapServer process is in normal state.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Go to the cluster details page and choose **Alarms**.
2. Check whether ALM-12007 Process Fault is generated.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
3. Check whether the service name and host name in the alarm are consistent with the LdapServer service and host names.
   - If yes, go to **Step 2.4**.
   - If no, go to **Step 3**.
4. Rectify the fault by following steps provided in ALM-12007 Process Fault.
5. In the alarm list, check whether the alarm ALM-25000 LdapServer Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.529 ALM-25004 Abnormal LdapServer Data Synchronization (For MRS 2.*x* or Earlier)

Description

This alarm is generated when LdapServer data on Manager is inconsistent. This alarm is cleared when the data becomes consistent.

This alarm is generated when LdapServer data in the cluster is inconsistent with LdapServer data on Manager. This alarm is cleared when the data becomes consistent.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25004	Critical	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

LdapServer data inconsistency occurs because LdapServer data on Manager or in the cluster is damaged. The LdapServer process with damaged data cannot provide services externally, and the authentication functions of Manager and the cluster are affected.

Possible Causes

- The network of the node where the LdapServer process locates is faulty.
- The LdapServer process is abnormal.
- The OS restart damages data on LdapServer.

Procedure

Step 1 Check whether the network where the LdapServer nodes reside is faulty.

1. Go to the cluster details page and choose **Alarms**.
2. Record the IP address of **HostName** in **Location** of the alarm as **IP1** (if multiple alarms exist, record the IP addresses as **IP1**, **IP2**, and **IP3** respectively).
3. Contact O&M personnel and use PuTTY to log in to the node corresponding to **IP1**. Run the **ping** command on the node to check whether the IP address of the management plane of the active OMS node can be pinged.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2.1**.
4. Contact O&M personnel to recover the network and check whether the alarm **ALM-25004 Abnormal LdapServer Data Synchronization** is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check whether the LdapServer process is in normal state.

1. Go to the cluster details page and choose **Alarms**.
2. Check whether ALM-12004 OLdap Resource Is Abnormal is generated for LdapServer.
   - If yes, go to **Step 2.3**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If no, go to **Step 2.5**.

3. Rectify the fault by following steps provided in **ALM-12004 OLdap Resource Abnormal (For MRS 2.x or Earlier)**.
4. Check whether the alarm ALM-25004 Abnormal LdapServer Data Synchronization is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.5**.
5. On the **Alarms** page of MRS Manager, check whether the alarm ALM-12007 Process Fault of LdapServer is generated.
   - If yes, go to **Step 2.6**.
   - If no, go to **Step 3.1**.
6. Rectify the fault by following steps provided in **ALM-12007 Process Fault (For MRS 2.x or Earlier)**.
7. Check whether the alarm ALM-25004 Abnormal LdapServer Data Synchronization is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3.1**.

Step 3 Check whether the OS restart damages data on LdapServer.

1. Go to the cluster details page and choose **Alarms**.
2. Record the IP address of **HostName** in **Location** of the alarm as **IP1** (if multiple alarms exist, record the IP addresses as **IP1**, **IP2**, and **IP3** respectively). Choose **Services** > **LdapServer** > **Service Configuration** and record the LdapServer port number as **PORT**. (If the IP address in the alarm location information is the IP address of the standby OMS node, the default port number is 21750.)
3. Log in to node **IP1** as user **omm** and run the **ldapsearch -H ldaps://IP1:PORT -x -LLL -b dc=hadoop,dc=com** command (if the IP address is the IP address of the standby OMS node, run the **ldapsearch -H ldaps://IP1:PORT -x -LLL -b dc=hadoop,dc=com** command before running this command). Check whether error information is displayed in the command output.
   - If yes, go to **Step 3.4**.
   - If no, go to **Step 4**.
4. Use the backup package generated before the alarm generation date to restore the LdapServer and OMS.

   NOTE

   Use the OMS data and LdapServer data backed up at the same time to restore data. Otherwise, the service and operation may fail. To recover data when services run properly, you are advised to manually back up the latest management data and then recover the data. Otherwise, Manager data produced between the backup point in time and the recovery point in time will be lost.
5. Check whether the alarm ALM-25004 Abnormal LdapServer Data Synchronization is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Reference

None

9.11.530 ALM-25500 KrbServer Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the KrbServer service status every 30 seconds. This alarm is generated when the KrbServer service is abnormal.

This alarm is cleared when the KrbServer service is in normal state.

Attribute

Alarm ID	Alarm Severity	Auto Clear
25500	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

When this alarm is generated, no operation can be performed for the KrbServer component in the cluster. The authentication of KrbServer in other components will be affected. The health status of components that depend on KrbServer in the cluster is **Bad**.

Possible Causes

- The node where the KrbServer service locates is faulty.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The OLdap service is unavailable.

Procedure

Step 1 Check whether the node where the KrbServer service locates is faulty.

1. Go to the MRS cluster details page and click **Components**.
2. Choose **KrbServer** > **Instances**. Go to the KrbServer instance page and view the host name of the node where the KrbServer service is deployed.
3. On the **Alarms** page of MRS Manager, check whether the alarm ALM-12006 Node Fault is generated.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2.1**.
4. Check whether the host name in the alarm information is the same as the actual host name in **Step 1.2**.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2.1**.
5. Rectify the fault by following steps provided in ALM-12006 Node Fault.
6. In the alarm list, check whether the alarm ALM-25500 KrbServer Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 2 Check whether the OLdap service is unavailable.

1. Go to the cluster details page and choose **Alarms**.
2. Check whether ALM-12004 OLdap Resource Is Abnormal is generated.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 3**.
3. Rectify the fault by following steps provided in ALM-12004 OLdap Resource Is Abnormal.
4. In the alarm list, check whether the alarm ALM-25500 KrbServer Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 3**.

Step 3 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.531 ALM-26051 Storm Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Storm service availability every 30 seconds. This alarm is generated if the Storm service becomes unavailable after all Nimbus nodes in a cluster become abnormal.

This alarm is cleared after the Storm service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
26051	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

- The cluster cannot provide the Storm service.
- Users cannot run new Storm tasks.

Possible Causes

- The Kerberos component is faulty.
- ZooKeeper is faulty or suspended.
- The active and standby Nimbus nodes in the Storm cluster are abnormal.

Procedure

Step 1 Check the Kerberos component status. For clusters without Kerberos authentication, skip this step and go to **Step 2**.

1. Go to the MRS cluster details page and click **Components**.
2. Check whether the health status of the Kerberos service is **Good**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 2.1**.
- If no, go to **Step 1.3**.

3. Rectify the fault by following instructions in **ALM-25500 KrbServer Service Unavailable (For MRS 2.x or Earlier)**.
4. Perform **Step 1.2** again.

Step 2 Check the ZooKeeper component status.

1. Check whether the health status of the ZooKeeper service is **Good**.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 2.2**.
2. If the ZooKeeper service is stopped, start it. For other problems, follow the instructions in **ALM-13000 ZooKeeper Service Unavailable (For MRS 2.x or Earlier)**.
3. Perform **Step 2.1** again.

Step 3 Check the status of the active and standby Nimbus nodes.

1. Choose **Components** > **Storm** > **Nimbus**.
2. In **Role**, check whether only one active Nimbus node exists.
   - If yes, go to **Step 4**.
   - If no, go to **Step 3.3**.
3. Select the two Nimbus instances and choose **More** > **Restart Instance**. Check whether the restart is successful.
   - If yes, go to **Step 3.4**.
   - If no, go to **Step 4**.
4. Log in to MRS Manager again and choose **Components** > **Storm** > **Nimbus**. Check whether the health status of Nimbus is **Good**.
   - If yes, go to **Step 3.5**.
   - If no, go to **Step 4**.
5. Wait 30 seconds and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

N/A

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.532 ALM-26052 Number of Available Supervisors in Storm Is Lower Than the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the number of supervisors every 60 seconds and compares it with the threshold. This alarm is generated if the number of supervisors is lower than the threshold.

To modify the threshold, users can choose **System** > **Threshold Configuration** on MRS Manager.

This alarm is cleared if the number of supervisors is greater than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
26052	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

- Existing tasks in the cluster cannot be executed.
- The cluster can receive new Storm tasks but cannot execute them.

Possible Causes

Supervisors are abnormal in the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Check the supervisor status.

1. Go to the cluster details page and click **Components**.
2. Choose **Storm** > **Supervisor**.
3. In **Role**, check whether the cluster has supervisor instances that are in the **Faulty** or **Recovering** state.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2**.
4. Select the supervisor instances that are in the **Faulty** or **Recovering** state and choose **More** > **Restart Instance**.
   - If yes, go to **Step 1.5**.
   - If the restart fails, go to **Step 2**.
5. Wait 30 seconds and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.533 ALM-26053 Slot Usage of Storm Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the slot usage of Storm every 60 seconds and compares it with the threshold. This alarm is generated if the slot usage exceeds the threshold.

To modify the threshold, users can choose **System** > **Threshold Configuration** on MRS Manager.

This alarm is cleared if the slot usage is lower than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
26053	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Users cannot run new Storm tasks.

Possible Causes

- Supervisors are abnormal in the cluster.
- Supervisors are normal but have poor processing capability.

Procedure

Step 1 Check the supervisor status.

1. Go to the cluster details page and click **Components**.
2. Choose **Storm** > **Supervisor**.
3. In **Role**, check whether the cluster has supervisor instances that are in the **Faulty** or **Recovering** state.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2.1** or **Step 3.1**.
4. Select the supervisor instances that are in the **Faulty** or **Recovering** state and choose **More** > **Restart Instance**.
   - If yes, go to **Step 1.5**.
   - If the restart fails, go to **Step 4**.
5. Wait a moment and then check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2.1** or **Step 3.1**.

Step 2 Increase the number of slots for the supervisors.

1. Go to the cluster details page and click **Components**.
2. Choose **Storm** > **Supervisor** > **Service Configuration**, and set **Type** to **All**.
3. Increase the value of **supervisor.slots.ports** to increase the number of slots for each supervisor. Then restart the instances.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Wait a moment and then check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 3 Expand the capacity of the supervisors.

1. Add nodes.
2. Wait a moment and then check whether the alarm is cleared.
   - If yes, no further action is required.
   - If the restart fails, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.534 ALM-26054 Heap Memory Usage of Storm Nimbus Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the heap memory usage of Storm Nimbus every 30 seconds and compares it with the threshold. This alarm is generated if the heap memory usage exceeds the threshold (80% by default).

To modify the threshold, users can choose **System** > **Threshold Configuration** > **Service** > **Storm** on MRS Manager.

This alarm is cleared if the heap memory usage is lower than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
26054	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Frequent memory garbage collection or memory overflow may occur, affecting submission of Storm services.

Possible Causes

The heap memory usage is high or the heap memory is improperly allocated.

Procedure

Step 1 Check the heap memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Choose **ALM-26054 Heap Memory Usage of Storm Nimbus Exceeds the Threshold > Location**. Query the **HostName** of the alarmed instance.
3. Choose **Components > Storm > Instances > Nimbus (corresponding to the HostName of the alarmed instance) > Customize > Heap Memory Usage of Nimbus**.
4. Check whether the heap memory usage of Nimbus has reached the threshold (80%).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Adjust the heap memory.

   Choose **Components > Storm > Service Configuration**, and set **Type** to **All**. Choose **Nimbus** > **System**. Increase the value of **-Xmx** in **NIMBUS_GC_OPTS**. Click **Save Configuration**. Select **Restart the affected services or instances** and click **OK**.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Related Information

N/A

9.11.535 ALM-27001 DBService Unavailable (For MRS 2.*x* or Earlier)

Description

The alarm module checks the DBService status every 30 seconds. This alarm is generated when the system detects that DBService is unavailable.

This alarm is cleared when DBService recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
27001	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The database service is unavailable and cannot provide data import and query functions for upper-layer services, which results in service exceptions.

Possible Causes

- The floating IP address does not exist.
- There is no active DBServer instance.
- The active and standby DBServer processes are abnormal.

Procedure

Step 1 Check whether the floating IP address exists in the cluster environment.

1. Go to the MRS cluster details page and click **Components**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Choose **DBService** > **Instances**.
3. Check whether the active instance exists.
   - If yes, go to **Step 1.4**.
   - If no, go to **Step 2.1**.
4. Select the active DBServer instance and record the IP address.
5. Log in to the host with the preceding IP address and run the **ifconfig** command to check whether the DBService floating IP address exists on the node.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 2.1**.
6. Run the **ping** *floating IP address* command to check whether the DBService floating IP address can be pinged.
   - If yes, go to **Step 1.7**.
   - If no, go to **Step 2.1**.
7. Log in to the host where the DBService floating IP address is located and run the **ifconfig** *interface* **down** command to delete the floating IP address.
8. Choose **Components** > **DBService** > **More** > **Restart Service** to restart DBService and check whether DBService is started successfully.
   - If yes, go to **Step 1.9**.
   - If no, go to **Step 2.1**.
9. Wait about 2 minutes and check whether the alarm is cleared in the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 13**.

Step 2 Check the status of the active DBServer instance.

1. Select the DBServer instance whose role status is abnormal and record the IP address.
2. On the **Alarms** page, check whether ALM-12007 Process Fault occurs in the DBServer instance on the host that corresponds to the IP address.
   - If yes, go to **Step 2.3**.
   - If no, go to **Step 4**.
3. Rectify the fault by following steps provided in ALM-12007 Process Fault.
4. Wait about 5 minutes and check whether the alarm is cleared in the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 3 Check the status of the active and standby DBServers.

1. Log in to the host where the DBService floating IP address is located, run the **sudo su - root** and **su - omm** commands to switch to user **omm**, and run the **cd ${BIGDATA_HOME}/FusionInsight/dbservice/** command to go to the DBService installation directory.
2. Run the **sh sbin/status-dbserver.sh** command to view the status of the active and standby HA processes of DBService. Determine whether the status can be viewed successfully.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If yes, go to **Step 3.3**.
- If no, go to **Step 4**.

3. Check whether the active and standby HA processes are abnormal.
   - If yes, go to **Step 3.4**.
   - If no, go to **Step 4**.
4. Choose **Components** > **DBService** > **More** > **Restart Service** to restart DBService and check whether DBService is started successfully.
   - If yes, go to **Step 3.5**.
   - If no, go to **Step 4**.
5. Wait about 2 minutes and check whether the alarm is cleared in the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.536 ALM-27003 DBService Heartbeat Interruption Between the Active and Standby Nodes (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the active or standby DBService node does not receive heartbeat messages from the peer node.

This alarm is cleared when the heartbeat recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
27003	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Local DBService HA Name	Specifies a local DBService HA.
Peer DBService HA Name	Specifies a peer DBService HA.

Impact on the System

During the DBService heartbeat interruption, only one node can provide the service. If this node is faulty, no standby node is available for failover and the service is unavailable.

Possible Causes

The link between the active and standby DBService nodes is abnormal.

Procedure

Step 1 Check whether the network between the active and standby DBService servers is in normal state.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, locate the row that contains the alarm and view the IP address of the standby DBService server in the alarm details.
3. Log in to the active DBService server.
4. Run the **ping** *heartbeat IP address of the standby DBService* command to check whether the standby DBService server is reachable.
   - If yes, go to **Step 2**.
   - If no, go to **Step 1.5**.
5. Contact the network administrator to check whether the network is faulty.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 2**.
6. Rectify the network fault and check whether the alarm is cleared from the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.537 ALM-27004 Data Inconsistency Between Active and Standby DBServices (For MRS 2.*x* or Earlier)

Description

The system checks the data synchronization status between the active and standby DBServices every 10 seconds. This alarm is generated when the synchronization status cannot be queried for six consecutive times or when the synchronization status is abnormal.

This alarm is cleared when the synchronization is in normal state.

Attribute

Alarm ID	Alarm Severity	Auto Clear
27004	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Local DBService HA Name	Specifies a local DBService HA.
Peer DBService HA Name	Specifies a peer DBService HA.
SYNC_PERSENT	Synchronization percentage.

Impact on the System

When data is not synchronized between the active and standby DBServices, the data may be lost or abnormal if the active instance becomes abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The network between the active and standby nodes is unstable.
- The standby DBService is abnormal.
- The disk space of the standby node is full.

Procedure

Step 1 Check whether the network between the active and standby nodes is in normal state.

1. Go to the cluster details page and choose **Alarms**.
2. In the alarm list, locate the row that contains the alarm and view the IP address of the standby DBService node in the alarm details.
3. Log in to the active DBService node.
4. Run the **ping** *heartbeat IP address of the standby DBService* command to check whether the standby DBService node is reachable.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.5**.
5. Contact the O&M personnel to check whether the network is faulty.
   - If yes, go to **Step 1.6**.
   - If no, go to **Step 2.1**.
6. Rectify the network fault and check whether the alarm is cleared from the alarm list.
   - If yes, no further action is required.
   - If no, go to **Step 2.1**.

Step 2 Check whether the standby DBService is in normal state.

1. Log in to the standby DBService node.
2. Run the following commands to switch the user:

   sudo su - root

   su - omm
3. Go to the **${DBSERVER_HOME}/sbin** directory and run the **./status-dbserver.sh** command to check whether the GaussDB resource status of the standby DBService is in normal state. In the command output, check whether the following information is displayed in the row where **ResName** is **gaussDB**:

   Example:

   ```
   10_10_10_231 gaussDB Standby_normal Normal Active_standby
   ```
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 4**.

Step 3 Check whether the disk space of the standby node is insufficient.

1. Log in to the standby DBService node.
2. Run the following commands to switch the user:

   sudo su - root

   su - omm

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Go to the **${DBSERVER_HOME}** directory, and run the following commands to obtain the DBService data directory:

   cd ${DBSERVER_HOME}

   source .dbservice_profile

   echo ${DBSERVICE_DATA_DIR}
4. Run the **df -h** command to check the system disk partition usage.
5. Check whether the DBService data directory space is full.
   - If yes, go to **Step 3.6**.
   - If no, go to **Step 4**.
6. Perform upgrade and expand capacity.
7. After capacity expansion, wait 2 minutes and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.538 ALM-28001 Spark Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Spark service status every 30 seconds. This alarm is generated when the Spark service is unavailable.

This alarm is cleared when the Spark service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
28001	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The Spark tasks submitted by users fail to be executed.

Possible Causes

- The KrbServer service is abnormal.
- The LdapServer service is abnormal.
- The ZooKeeper service is abnormal.
- The HDFS service is abnormal.
- The Yarn service is abnormal.
- The corresponding Hive service is abnormal.

Procedure

Step 1 Check whether service unavailability alarms exist in services that Spark depends on.

1. Go to the MRS cluster details page and choose **Alarms**.
2. Check whether the following alarms exist in the alarm list:
   a. ALM-25500 KrbServer Service Unavailable
   b. ALM-25000 LdapServer Service Unavailable
   c. ALM-13000 ZooKeeper Service Unavailable
   d. ALM-14000 HDFS Service Unavailable
   e. ALM-18000 Yarn Service Unavailable
   f. ALM-16004 Hive Service Unavailable
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 2**.
3. Handle the alarms based on the troubleshooting methods provided in the alarm help.

   After the alarm is cleared, wait a few minutes and check whether the alarm HetuServer Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.539 ALM-38000 Kafka Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Kafka service availability every 30 seconds. This alarm is generated if the Kafka service becomes unavailable.

This alarm is cleared after the Kafka service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
38000	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The cluster cannot provide the Kafka service and users cannot run new Kafka tasks.

Possible Causes

- The KrbServer component is faulty.
- The ZooKeeper component is faulty or fails to respond.
- The Broker node in the Kafka cluster is abnormal.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Check the KrbServer component status. For clusters without Kerberos authentication, skip this step and go to **Step 2**.

1. Go to the MRS cluster details page and click **Components**.
2. Check whether the health status of the KrbServer service is **Good**.
   - If yes, go to **Step 2.1**.
   - If no, go to **Step 1.3**.
3. Rectify the fault by following instructions in **ALM-25500 KrbServer Service Unavailable (For MRS 2.x or Earlier)**.
4. Perform **Step 1.2** again.

Step 2 Check the ZooKeeper component status.

1. Check whether the health status of the ZooKeeper service is **Good**.
   - If yes, go to **Step 3.1**.
   - If no, go to **Step 2.2**.
2. If the ZooKeeper service is stopped, start it. For other problems, follow the instructions in **ALM-13000 ZooKeeper Service Unavailable (For MRS 2.x or Earlier)**.
3. Perform **Step 2.1** again.

Step 3 Check the Broker status.

1. Choose **Components** > **Kafka** > **Broker**.
2. In **Role**, check whether all instances are normal.
   - If yes, go to **Step 3.4**.
   - If no, go to **Step 3.3**.
3. Select all instances of Broker and choose **More** > **Restart Instance**.
   - If the restart is successful, go to **Step 3.4**.
   - If the restart fails, go to **Step 4**.
4. Choose **Components > Kafka**. Check whether the health status of Kafka is **Good**.
   - If yes, go to **Step 3.5**.
   - If no, go to **Step 4**.
5. Wait 30 seconds and check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 4**.

Step 4 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

N/A

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.540 ALM-38001 Insufficient Kafka Disk Capacity (For MRS 2.*x* or Earlier)

Description

The system checks the Kafka disk usage every 60 seconds and compares it with the threshold. This alarm is generated if the disk usage exceeds the threshold.

To modify the threshold, users can choose **System** > **Threshold Configuration** on MRS Manager.

This alarm is cleared if the Kafka disk usage is lower than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
38001	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
PartitionName	Specifies the disk partition where the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Kafka fails to write data to the disks.

Possible Causes

- The Kafka disk configurations (such as disk count and disk size) are insufficient for the data volume.
- The data retention period is long and historical data occupies large space.
- Services are improperly planned. As a result, data is unevenly distributed and some disks are full.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Step 1 Go to the MRS cluster details page and choose **Alarms**.

Step 2 In the alarm list, click the alarm and view the **HostName** and **PartitionName** of the alarm in **Location** of **Alarm Details**.

Step 3 On the **Hosts** page, click the host name obtained in **Step 2**.

Step 4 Check whether the **Disk** area contains the **PartitionName** of the alarm.

- If yes, go to **Step 5**.
- If no, manually clear the alarm and no further action is required.

Step 5 In the **Disk** area, check whether the usage of the alarmed partition has reached 100%.

- If yes, go to **Step 6**.
- If no, go to **Step 8**.

Step 6 In **Instance**, choose **Broker > Instance Configuration**. On the **Instance Configuration** page that is displayed, set **Type** to **All** and query the data directory parameter **log.dirs**.

Step 7 Choose **Components > Kafka > Instances**. On the **Kafka Instance** page that is displayed, stop the Broker instance corresponding to **Step 2**. Then log in to the alarmed node and manually delete the data directory in **Step 6**. After all subsequent operations are complete, start the Broker instance.

Step 8 Choose **Components > Kafka > Service Configuration**. The **Kafka Configuration** page is displayed.

Step 9 Check whether **disk.adapter.enable** is **true**.

- If yes, go to **Step 11**.
- If no, change the value to **true** and go to **Step 10**.

Step 10 Check whether the **adapter.topic.min.retention.hours** parameter, indicating the minimum data retention period, is properly configured.

- If yes, go to **Step 12**.
- If no, set it to a proper value and go to **Step 12**.

NOTE

If the retention period cannot be adjusted for certain topics, the topics can be added to **disk.adapter.topic.blacklist**.

Step 11 Wait 10 minutes and check whether the disk usage is reduced.

- If yes, wait until the alarm is cleared.
- If no, go to **Step 12**.

Step 12 Go to the **Kafka Topic Monitor** page and query the data retention period configured for Kafka. Determine whether the retention period needs to be shortened based on service requirements and data volume.

- If yes, go to **Step 13**.
- If no, go to **Step 14**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 13 Find the topics with great data volumes based on the disk partition obtained in **Step 2**. Log in to the Kafka client and manually shorten the data retention period for these topics using the following command:

kafka-topics.sh --zookeeper *ZooKeeper address:24002/kafka* **--alter --topic** *Topic name* **--config retention.ms=***Retention period*

Step 14 Check whether partitions are properly configured for topics. For example, if the number of partitions for a topic with a large data volume is smaller than the number of disks, data may be unevenly distributed to the disks and the usage of some disks will reach the upper limit.

NOTE

To identify topics with large data volumes, log in to the relevant nodes obtained in **Step 2**, go to the data directory (the configured directory before **log.dirs** in **Step 6** is modified), and check the disk space used by each topic's partitions.

- If the partitions are improperly configured, go to **Step 15**.
- If the partitions are properly configured, go to **Step 16**.

Step 15 On the Kafka client, add partitions to the topics.

kafka-topics.sh --zookeeper *ZooKeeper address:24002/kafka* **--alter --topic** *Topic name* **--partitions=***Number of new partitions*

NOTE

It is advised to set the number of new partitions to a multiple of the number of Kafka disks.

This operation may not quickly clear the alarm. Data will be gradually balanced among the disks.

Step 16 Check whether the cluster capacity needs to be expanded.

- If yes, add nodes to the cluster and go to **Step 17**.
- If no, go to **Step 17**.

Step 17 Wait a moment and then check whether the alarm is cleared.

- If yes, no further action is required.
- If no, go to **Step 18**.

Step 18 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.541 ALM-38002 Heap Memory Usage of Kafka Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the heap memory usage of Kafka every 30 seconds. This alarm is generated if the heap memory usage of Kafka exceeds the threshold (80%).

This alarm is cleared if the heap memory usage is lower than the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
38002	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.
Trigger Condition	Generates an alarm when the actual indicator value exceeds the specified threshold.

Impact on the System

Memory overflow may occur, causing service crashes.

Possible Causes

The heap memory usage is high or the heap memory is improperly allocated.

Procedure

Step 1 Check the heap memory usage.

1. Go to the MRS cluster details page and choose **Alarms**.
2. Choose **ALM-38002 Kafka Heap Memory Usage Exceeds the Threshold** > **Location**. Query the IP address of the alarmed instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Choose **Components > Kafka > Instance > Broker (corresponding to the IP address of the alarmed instance) > Customize > Kafka Heap Memory Resource Percentage** to check the heap memory usage.
4. Check whether the heap memory usage of Kafka has reached the threshold (80%).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Kafka** > **Service Configuration** > **All** > **Broker** > **Environment Variables**. Increase the value of **KAFKA_HEAP_OPTS** as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Related Information

N/A

9.11.542 ALM-43001 Spark Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Spark service status every 60 seconds. This alarm is generated when the Spark service is unavailable.

This alarm is cleared when the Spark service recovers.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43001	Critical	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

The Spark tasks submitted by users fail to be executed.

Possible Causes

- The KrbServer service is abnormal.
- The LdapServer service is abnormal.
- ZooKeeper is abnormal.
- The HDFS service is abnormal.
- The Yarn service is abnormal.
- The corresponding Hive service is abnormal.

Procedure

Step 1 Check whether service unavailability alarms exist in services that Spark depends on.

1. Go to the cluster details page and choose **Alarms**.
2. Check whether the following alarms exist in the alarm list:
   a. ALM-25500 KrbServer Service Unavailable
   b. ALM-25000 LdapServer Service Unavailable
   c. ALM-13000 ZooKeeper Service Unavailable
   d. ALM-14000 HDFS Service Unavailable
   e. ALM-18000 Yarn Service Unavailable
   f. ALM-16004 Hive Service Unavailable
   - If yes, go to **Step 1.3**.
   - If no, go to **Step 2**.
3. Handle the alarm according to the alarm help.
   After the alarm is cleared, wait a few minutes and check whether the alarm HetuServer Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.543 ALM-43006 Heap Memory Usage of the JobHistory Process Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the JobHistory process status every 30 seconds. The alarm is generated when the heap memory usage of the JobHistory process exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43006	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the available JobHistory process heap memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The heap memory of the JobHistory process is overused or the heap memory is inappropriately allocated.

Procedure

Step 1 Check the heap memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43006** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Spark** > **Instance** > **JobHistory** (IP address of the instance for which the alarm is generated) > **Customize** > **Heap Memory Statistics of the JobHistory Process**. Click **OK** to view the heap memory usage.
4. Check whether the used heap memory of JobHistory reaches 90% of the maximum heap memory specified for JobHistory.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JobHistory** > **Default**. Increase the value of **SPARK_DAEMON_MEMORY** as required.
6. Click **Save Configuration** and select **Restart the affected services or instances**. Click **OK**.
7. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.544 ALM-43007 Non-Heap Memory Usage of the JobHistory Process Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the JobHistory process status every 30 seconds. The alarm is generated when the non-heap memory usage of the JobHistory process exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43007	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the available JobHistory process non-heap memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The non-heap memory of the JobHistory process is overused or the non-heap memory is inappropriately allocated.

Procedure

Step 1 Check non-heap memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43007** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Spark** > **Instance** > **JobHistory** (IP address of the instance for which the alarm is generated) > **Customize** > **Non-Heap Memory Statistics of the JobHistory Process**. Click **OK** to view the non-heap memory usage.
4. Check whether the non-heap memory usage of JobHistory has reached the threshold (90% of the maximum memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JobHistory** > **Default**. Increase the value of **-XX:MaxMetaspaceSize** in **SPARK_DAEMON_JAVA_OPTS** as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.545 ALM-43008 Direct Memory Usage of the JobHistory Process Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the JobHistory process status every 30 seconds. The alarm is generated when the direct memory usage of the JobHistory process exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43008	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the available JobHistory process direct memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The direct memory of the JobHistory process is overused or the direct memory is inappropriately allocated.

Procedure

Step 1 Check the direct memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43008** and view the IP address and role name of the instance in **Location**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Choose **Components** > **Spark** > **Instance** > **JobHistory** (IP address of the instance for which the alarm is generated) > **Customize** > **Direct Memory Statistics of the JobHistory Process**. Click **OK** to view the direct memory usage.
4. Check whether the direct memory usage of the JobHistory process has reached the threshold (90% of the maximum direct memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JobHistory** > **Default**. Increase the value of **-XX:MaxDirectMemorySize** in **SPARK_DAEMON_JAVA_OPTS** as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.546 ALM-43009 JobHistory GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the GC time of the JobHistory process every 60 seconds. This alarm is generated when the detected GC time exceeds the threshold (12 seconds) for three consecutive times. You can change the threshold by choosing **System** > **Threshold Configuration** > **Service** > **Spark** > **JobHistory GC Time** > **Total JobHistory GC Time**. This alarm is cleared when the JobHistory GC time is shorter than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43009	Major	Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the GC time exceeds the threshold, JobHistory may run in low performance.

Possible Causes

The heap memory of the JobHistory process is overused or inappropriately allocated, causing frequent GC.

Procedure

Step 1 Check the GC time.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43009** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Spark** > **Instance** > **JobHistory** (IP address of the instance for which the alarm is generated) > **Customize** > **GC Time of the JobHistory Process**. Click **OK** to view the GC time.
4. Check whether the GC time of the JobHistory process is longer than 12 seconds.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JobHistory** > **Default**. Increase the value of the **SPARK_DAEMON_MEMORY** parameter as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Reference

None

9.11.547 ALM-43010 Heap Memory Usage of the JDBCServer Process Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the JDBCServer process status every 30 seconds. The alarm is generated when the heap memory usage of the JDBCServer process exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43010	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the available JDBCServer process heap memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The heap memory of the JDBCServer process is overused or the heap memory is inappropriately allocated.

Procedure

Step 1 Check the heap memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43010** and view the IP address and role name of the instance in **Location**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Choose **Components** > **Spark** > **Instance** > **JDBCServer** (IP address of the instance for which the alarm is generated) > **Customize** > **Heap Memory Statistics of the JDBCServer Process**. Click **OK** to view the heap memory usage.
4. Check whether the heap memory usage of JDBCServer has reached the threshold (90% of the maximum heap memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JDBCServer** > **Tuning**. Increase the value of the **SPARK_DRIVER_MEMORY** parameter as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.548 ALM-43011 Non-Heap Memory Usage of the JDBCServer Process Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the JDBCServer process status every 30 seconds. The alarm is generated when the non-heap memory usage of the JDBCServer process exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43011	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the available JDBCServer process non-heap memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The non-heap memory of the JDBCServer process is overused or the non-heap memory is inappropriately allocated.

Procedure

Step 1 Check non-heap memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43011** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Spark** > **Instance** > **JDBCServer** (IP address of the instance for which the alarm is generated) > **Customize** > **Non-heap Memory Statistics of the JDBCServer Process**. Click **OK** to view the non-heap memory usage.
4. Check whether the non-heap memory usage of JDBCServer has reached the threshold (90% of the maximum non-heap memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JDBCServer** > **Tuning**. Increase the value of **-XX:MaxMetaspaceSize** in **spark.driver.extraJavaOptions** as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.549 ALM-43012 Direct Memory Usage of the JDBCServer Process Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the JDBCServer process status every 30 seconds. The alarm is generated when the direct memory usage of the JDBCServer process exceeds the threshold (90% of the maximum memory).

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43012	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the available JDBCServer process direct memory is insufficient, a memory overflow occurs and the service breaks down.

Possible Causes

The direct memory of the JDBCServer process is overused or the direct memory is inappropriately allocated.

Procedure

Step 1 Check the direct memory usage.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43012** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Spark** > **Instance** > **JDBCServer** (IP address of the instance for which the alarm is generated) > **Customize** > **Direct Memory Statistics of the JDBCServer Process**. Click **OK** to view the direct memory usage.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Check whether the direct memory usage of the JDBCServer process has reached the threshold (90% of the maximum direct memory).
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JDBCServer** > **Tuning**. Increase the value of **-XX:MaxDirectMemorySize** in **spark.driver.extraJavaOptions** as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.550 ALM-43013 JDBCServer GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system checks the GC time of the JDBCServer process every 60 seconds. This alarm is generated when the detected GC time exceeds the threshold (12 seconds) for three consecutive times. You can change the threshold by choosing **System** > **Threshold Configuration** > **Service** > **Spark** > **JDBCServer GC Time** > **Total JDBCServer GC Time**. This alarm is cleared when the JDBCServer GC time is shorter than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Automatically Cleared
43013	Major	Yes

Parameters

Parameter	Description
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
HostName	Specifies the host for which the alarm is generated.

Impact on the System

If the GC time exceeds the threshold, JDBCServer may run in low performance.

Possible Causes

The heap memory of the JDBCServer process is overused or inappropriately allocated, causing frequent GC.

Procedure

Step 1 Check the GC time.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **43013** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Spark** > **Instance** > **JDBCServer** (IP address of the instance for which the alarm is generated) > **Customize** > **GC Time of the JDBCServer Process**. Click **OK** to view the GC time.
4. Check whether the GC time of the JDBCServer process is longer than 12 seconds.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Spark** > **Service Configuration**. Set **Type** to **All** and choose **JDBCServer** > **Tuning**. Increase the value of the **SPARK_DRIVER_MEMORY** parameter as required.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

9.11.551 ALM-44004 Presto Coordinator Resource Group Queuing Tasks Exceed the Threshold (For MRS 2.*x* or Earlier)

Description

This alarm is generated when the system detects that the number of queuing tasks in a resource group exceeds the threshold. The system queries the number of queuing tasks in a resource group through the JMX interface. You can choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Presto** > **resource-groups** to configure a resource group. You can choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Coordinator** > **Customize** > **resourceGroupAlarm** to configure the threshold of each resource group.

Attribute

Alarm ID	Alarm Severity	Auto Clear
44004	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

If the number of queuing tasks in a resource group exceeds the threshold, a large number of tasks may be in the queuing state. The Presto task time exceeds the expected value. When the number of queuing tasks in a resource group exceeds the maximum number (**maxQueued**) of queuing tasks in the resource group, new tasks cannot be executed.

Possible Causes

The resource group configuration is improper or too many tasks in the resource group are submitted.

Procedure

Step 1 Choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Presto** > **resource-groups** to adjust the resource group configuration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 You can choose **Components** > **Presto** > **Service Configuration** (switch **Basic** to **All**) > **Coordinator** > **Customize** > **resourceGroupAlarm** to modify the threshold of each resource group.

Step 3 Collect fault information.

1. Log in to the cluster node based on the host name in the fault information and query the number of queuing tasks on the Presto client based on **Resource Group** in the additional information.
2. Log in to the cluster node based on the host name in the fault information, view the **/var/log/Bigdata/nodeagent/monitorlog/monitor.log** file, and search for resource group information to view the monitoring collection information of the resource group.
3. Contact O&M engineers and send the collected logs.

----End

Related Information

None

9.11.552 ALM-44005 Presto Coordinator Process GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system collects GC time of the Presto Coordinator process every 30 seconds. This alarm is generated when the GC time exceeds the threshold (exceeds 5 seconds for three consecutive times). You can change the threshold by choosing **System** > **Configure Alarm Threshold** > **Service** > **Presto** > **Coordinator** > **Presto Process Garbage Collection Time** > **Garbage Collection Time of the Coordinator Process** on MRS Manager. This alarm is cleared when the Coordinator process GC time is less than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
44005	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Impact on the System

If the GC time of the Coordinator process is too long, the Coordinator process running performance will be affected and the Coordinator process will even be unavailable.

Possible Causes

The heap memory of the Coordinator process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Procedure

Step 1 Check the GC time.

1. Go to the MRS cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **44005** and view the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Presto** > **Instances** > **Coordinator** (business IP address of the instance for which the alarm is generated) > **Customize** > **Presto Garbage Collection Time**. Click **OK** to view the GC time.
4. Check whether the GC time of the Coordinator process is longer than 5 seconds.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Presto** > **Service Configuration**, and switch **Basic** to **All**. Choose **Presto** > **Coordinator**. Increase the value of **-Xmx** (maximum heap memory) in the **JAVA_OPTS** parameter based on the site requirements.
6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact O&M engineers and send the collected logs.

----End

Related Information

None

9.11.553 ALM-44006 Presto Worker Process GC Time Exceeds the Threshold (For MRS 2.*x* or Earlier)

Description

The system collects GC time of the Presto Worker process every 30 seconds. This alarm is generated when the GC time exceeds the threshold (exceeds 5 seconds for three consecutive times). You can change the threshold by choosing **System** >

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configure Alarm Threshold > **Service** > **Presto** > **Worker** > **Presto Garbage Collection Time** > **Garbage Collection Time of the Worker Process** on MRS Manager. This alarm is cleared when the Worker process GC time is shorter than or equal to the threshold.

Attribute

Alarm ID	Alarm Severity	Auto Clear
44006	Major	Yes

Parameter

Parameter	Description
ServiceName	Service for which the alarm is generated.
RoleName	Role for which the alarm is generated.
HostName	Host for which the alarm is generated.

Impact on the System

If the GC time of the Worker process is too long, the Worker process running performance will be affected and the Worker process will even be unavailable.

Possible Causes

The heap memory of the Worker process is overused or inappropriately allocated, causing frequent occurrence of the GC process.

Procedure

Step 1 Check the GC time.

1. Go to the cluster details page and choose **Alarms**.
2. Select the alarm whose **Alarm ID** is **44006**. Then check the IP address and role name of the instance in **Location**.
3. Choose **Components** > **Presto** > **Instances** > **Worker** (business IP address of the instance for which the alarm is generated) > **Customize** > **Presto Garbage Collection Time**. Click **OK** to view the GC time.
4. Check whether the GC time of the Worker process is longer than 5 seconds.
   - If yes, go to **Step 1.5**.
   - If no, go to **Step 2**.
5. Choose **Components** > **Presto** > **Service Configuration**, and switch **Basic** to **All**, and choose **Presto** > **Worker** Increase the value of **-Xmx** (maximum heap memory) in the **JAVA_OPTS** parameter based on the site requirements.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

6. Check whether the alarm is cleared.
   - If yes, no further action is required.
   - If no, go to **Step 2**.

Step 2 Collect fault information.

1. On MRS Manager, choose **System** > **Export Log**.
2. Contact the O&M engineers and send the collected logs.

----End

Reference

None

9.11.554 ALM-45325 Presto Service Unavailable (For MRS 2.*x* or Earlier)

Description

The system checks the Presto service status every 60 seconds. This alarm is generated when the system detects that Presto is unavailable.

This alarm is cleared when the Presto service recovers.

Attribute

Alarm ID	Alarm Severity	Auto Clear
45325	Critical	Yes

Parameters

Name	Meaning
ServiceName	Specifies the service for which the alarm is generated.
RoleName	Specifies the role for which the alarm is generated.
HostName	Specifies the host for which the alarm is generated.

Impact on the System

Presto cannot run SQL queries.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Possible Causes

- The Presto coordinator or worker process is faulty.
- The network communication between Presto coordinator and worker instances is interrupted.

Procedure

Step 1 Check the status of the coordinator and worker processes.

1. Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Presto**. On the page that is displayed, click the **Instance** tab. In the Presto instance list, check whether the status of all coordinator or worker instances is **Unknown**.
   - If yes, go to **2**.
   - If no, go to **1**.
2. In the upper part of the Presto instance list, choose **More** > **Restart Service** to restart the coordinator and worker processes.
3. In the alarm list, check whether ALM-45325 Presto Service Unavailable is cleared.
   - If yes, no further action is required.
   - If no, go to **1** in **Step 2**.

Step 2 Collect fault information.

1. On FusionInsight Manager, choose **O&M**. In the navigation pane on the left, choose **Log** > **Download**.
2. Select **Presto** for **Service**.
3. Click  in the upper right corner, and set **Start Date** and **End Date** for log collection to 10 minutes ahead of and after the alarm generation time, respectively. Then, click **Download**.
4. Contact the O&M engineers and send the collected logs.

----End

Alarm Clearing

This alarm is automatically cleared after the fault is rectified.

9.12 Configuring Remote O&M for an MRS Cluster

If you encounter any issues while using a cluster and require assistance from support personnel, you can first contact them. Then, you can authorize them to access your machine through the O&M authorization function to locate the fault. Alternatively, you can provide support personnel with logs from a specific time period using the log sharing function to help them locate the fault.

Enabling Remote Node O&M Authorization for an MRS Cluster

This authorization will allow O&M personnel to temporarily access your cluster nodes before the expiration time. You can also manually cancel the authorization in advance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 1 Log in to the MRS console.

Step 2 In the navigation pane of the MRS console, choose **Active Clusters**, locate a running cluster, and click its name to switch to the cluster details page.

Step 3 In the upper right corner of the page, click **O&M**, choose **Authorize for Cluster Nodes**, and select the deadline for the support engineers to access the local host. Before the deadline, the support personnel have the temporary permission to access the local host.

Step 4 Select the check box to confirm authorization and click **OK**.

Step 5 After the fault is rectified, click **O&M** in the upper right corner of the page and choose **Deauthorize for Cluster Nodes** to retrieve the access permission granted to the support engineers.

----End

Configuring Remote O&M Log Sharing for an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 In the navigation pane of the MRS console, choose **Active Clusters**, locate a running cluster, and click its name to switch to the cluster details page.

Step 3 In the upper right corner of the displayed page, choose **O&M** > **Share Log**.

Step 4 In the displayed dialog box, select the start time and end time in **Time Range**.

- Select **Time Range** based on the suggestions of support personnel.
- **End Date** must be later than **Start Date**. Otherwise, logs cannot be filtered by time.

----End

Configuring Remote O&M Metric Sharing for an MRS Cluster

Step 1 Log in to the MRS console.

Step 2 In the navigation pane of the MRS console, choose **Active Clusters**, locate a running cluster, and click its name to switch to the cluster details page.

Step 3 In the upper right corner of the page, click **O&M** and select **Share Metric** to enable metric sharing.

Step 4 After the fault is rectified, click **O&M** in the upper right corner of the page and select **Disable Metric Sharing** to cancel the access permission for the support personnel.

----End

9.13 Common Ports for MRS Cluster Services

This section describes the common ports of each component in an MRS cluster. If the cluster version type is **LTS** during cluster creation in custom config mode as described in **Creating a Custom MRS Cluster**, you can determine whether to keep

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

the default ports of the components consistent with those of the open-source components.

- If **Component Port** is set to **Open source** or **Version Type** is set to **Normal**, find the default component ports in the **Default Open Source Port** column of the following table.
- If **Component Port** is set to **Custom**, find the default component ports in the **Default Custom Port** column of the following table.
- If there is only the **Default Port** column, the open-source ports of the component are the same as the default custom ports.
- The value range of each component port is a recommended value, defined by the product. However, the system does not enforce strict limitations on port ranges.

Common HBase Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
hbase.master.port	16000	21300	HMaster RPC port, which is used to connect the HBase client to HMaster. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
hbase.master.info.port	16010	21301	HMaster HTTPS port, which is used by the remote web client to connect to the HMaster UI. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
hbase.regionserver.port	16020	21302	RegionServer (RS) RPC port, which is used to connect the HBase client to RegionServer. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
hbase.regionserver.info.port	16030	21303	RegionServer HTTPS port, which is used by the remote web client to connect to the RegionServer UI. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
hbase.thrift.info.port	9095	21304	Thrift server listening port of ThriftServer. This port is used by client connections. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
hbase.regionserver.thrift.port	9090	21305	Thrift server listening port of RegionServer. This port is used by clients to connect to RegionServer. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
hbase.rest.info.port	8085	21308	Port of the native web UI of RegionServer RESTServer.
-	21309	21309	REST port of RegionServer RESTServer.

Common HDFS Ports

The protocol type of all ports in the table is TCP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
dfs.namenode.rpc.port	• 9820 (versions earlier than MRS 3.*x*) • 8020 (MRS 3.*x* or later)	25000	NameNode RPC port, which is used for communication between the HDFS client and NameNode and connections between different DataNodes. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
dfs.namenode.http.port	9870	25002	HTTP access port of NameNode. This port is used for the point-to-point NameNode checkpoint operations and for the remote web client to connect to the NameNode UI. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
dfs.namenode.https.port	9871	25003	HTTPS access port of NameNode. This port is used for the point-to-point NameNode checkpoint operations and for the remote web client to connect to the NameNode UI. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
dfs.datanode.ipc.port	9867	25008	Port for communication between DataNode processes. This port is used by clients to connect to DataNodes to perform RPC operations. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
dfs.datanode.port	9866	25009	DataNode data transmission port. This port is used by the HDFS client to transfer data from or to DataNode. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
dfs.datanode.http.port	9864	25010	DataNode HTTP port, which is used to remotely connect to the DataNode UI through a web client in secure mode. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
dfs.datanode.https.port	9865	25011	DataNode HTTPS port, which is used to remotely connect to the DataNode UI through a web client in secure mode. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
dfs.JournalNode.rpc.port	8485	25012	JournalNode RPC port, which is used for client communication. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
dfs.journalnode.http.port	8480	25013	JournalNode HTTP port, which is used to remotely connect to JournalNode through a web client in secure mode. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
dfs.journalnode.https.port	8481	25014	JournalNode HTTPS port, which is used to remotely connect to JournalNode through a web client in secure mode. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
httpfs.http.port	14000	25018	HttpFS HTTP server listening port, which is used to remotely connect to HttpFS through a REST API. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes

Common HetuEngine Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
server.port (HSBroker)	29860	29860	Specifies the port number that HSBroker listens to.
server.port (HSConsole)	29880	29880	Specifies the port number that HSConsole listens to.
server.port (HSFabric)	29900	29900	Specifies the port number that HSFabric listens to, which is used for cross-domain connections.
gateway.port	29902	29902	Specifies the port number that HSFabric listens to, which is used for JDBC connections.

Common Hive Ports

The protocol type of all ports in the table is TCP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
templeton.port	9111	21055	Port used for WebHCat to provide the REST service. This port is used for the communication between the WebHCat server and the WebHCat client. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
hive.server2.thrift.port	10000	21066	Port for HiveServer to provide Thrift services. This port is used for the communication between the HiveServer server and the HiveServer client. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
hive.metastore.port	9083	21088	Port for MetaStore to provide Thrift services. This port is used for communication between the Metastore client and Metastore, that is, communication between HiveServer and Metastore. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
hive.server2.webui.port	10002	-	Hive Web UI port, which is used to communicate with the Hive UI server using HTTPS/HTTP.

Common Hue Ports

The protocol type of all ports in the table is TCP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
HTTP_PORT	8888	21200	Port for Hue to provide HTTPS services This port is used to provide web services over HTTPS, which can be changed. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes

Common Kafka Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
port	9092	21005	Port for a broker to receive data and obtain services
ssl.port	9093	21008	SSL port used by a broker to receive data and obtain services
sasl.port	21007	21007	SASL security authentication port provided by a broker, which provides the secure Kafka service
sasl-ssl.port	21009	21009	Port used by a broker to provide encrypted service based on the SASL and SSL protocols

Common Loader Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Port	Port Description
LOADER_HTTPS_PORT	21351	This port is used to provide REST APIs for configuration and running of Loader jobs. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common Manager Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Port	Port Description
-	8080	Port provided by WebService for user access This port is used to access the web UI over HTTP. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
-	28443	Port provided by WebService for user access This port is used to access the web UI over HTTPS. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
-	-	CAS port This port is used for user login authentication. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
-	-	Control port provided by the backup and restoration plug-in process. This port is used by the controller to send backup and restoration control commands.
-	-	Port registered by the client with the OM server
-	-	OMS database service port This port is used to provide database services. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Port	Port Description
-	-	DBService database service port This port is used to provide database services. • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes
-	-	Port used to receive operation or query requests from the management plane and to forward the requests to the executor or manager for execution.

Common MapReduce Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
mapreduce.jobhistory.webapp.port	19888	26012	Web HTTP port of the JobHistory server This port is used to view the web page of the JobHistory server.
mapreduce.jobhistory.port	10020	26013	Port of the JobHistory server This port is used for: 1. Task data restoration in the MapReduce client 2. Obtaining task report in the Job client
mapreduce.jobhistory.webapp.https.port	19890	26014	Web HTTPS port of the JobHistory server This port is used to view the web page of the JobHistory server.

Common Spark Ports

The protocol type of all ports in the table is TCP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
hive.server2.thrift.port	22550	22550	JDBC thrift port This port is used for socket communication between the Spark 2.1.0 CLI/JDBC client and server. If **hive.server2.thrift.port** is used, an exception indicating that the port is occupied is reported. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
spark.ui.port	4040	22950	Web UI port of JDBC This port is used for HTTPS/HTTP communication between Web requests and the JDBC Server Web UI server. The system obtains the port based on the parameter settings and checks its validity. If the port is invalid, the system increases the port number by 1 each time for a maximum of 16 times until a valid port is obtained (the retry times can be configured using **spark.port.maxRetries**). ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
spark.history.ui.port	18080	22500	JobHistory Web UI port This port is used for HTTPS/HTTP communication between web requests and the Spark 2.1.0 History Server. The system obtains the port based on the parameter settings and checks its validity. If the port is invalid, the system increases the port number by 1 each time for a maximum of 16 times until a valid port is obtained (the retry times can be configured using **spark.port.maxRetries**). • Is the port enabled by default during the installation: Yes • Is the port enabled after security hardening: Yes

Common Storm Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
nimbus.thrift.port	6627	29200	Port for Nimbus to provide thrift services
supervisor.slots.ports	6700/6701/6702/6703	29200–29499	Port for receiving service requests that are forwarded from other servers
logviewer.https.port	29248	29248	Port for LogViewer to provide HTTPS services
ui.https.port	29243	29243	Port for the Storm UI to provide HTTPS services

Common YARN Ports

The protocol type of all ports in the table is TCP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
yarn.resourcemanager.webapp.port	8088	26000	Web HTTP port of the ResourceManager service
yarn.resourcemanager.webapp.https.port	8090	26001	Web HTTPS port of the ResourceManager service This port is used to access the Resource Manager web applications in security mode. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes
yarn.nodemanager.webapp.port	8042	26006	NodeManager Web HTTP port
yarn.nodemanager.webapp.https.port	8044	26010	NodeManager Web HTTPS port This port is used to access the NodeManager web applications in security mode. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes

Common ZooKeeper Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
clientPort	2181	24002	ZooKeeper client port This port is used for communication between the ZooKeeper client and ZooKeeper server. ● Is the port enabled by default during the installation: Yes ● Is the port enabled after security hardening: Yes

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common Kerberos Ports

The protocol type of all ports in the table is TCP and UDP.

Parameter	Default Port	Port Description
kdc_ports	21732	Kerberos server port. This port is used by components for Kerberos authentication. This port may be used when you configure mutual trust between clusters.

Common OpenTSDB Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Port	Port Description
tsd.network.port	4242	Web UI port of OpenTSDB This port is used for HTTPS/HTTP communication between web requests and the OpenTSDB UI server.

Common Tez Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Port	Port Description
tez.ui.port	28888	Web UI port of Tez

Common KafkaManager Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Port	Port Description
kafka_manager_port	9099	Web UI port of KafkaManager

Common Presto Ports

The protocol type of all ports in the table is TCP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Port	Port Description
http-server.http.port	7520	HTTP port for Presto coordinator to provide services to external systems
http-server.https.port	7521	HTTPS port for Presto coordinator to provide services to external systems
http-server.http.port	7530	HTTP port for Presto worker to provide services to external systems
http-server.https.port	7531	HTTPS port for Presto worker to provide services to external systems

Common Flink Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Port	Port Description
jobmanager.web.port	32261~32325	Web UI port of Flink This port is used for HTTP/HTTPS communication between the client web requests and Flink server.

Common ClickHouse Ports

The protocol type of the port in the table is TCP and HTTP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
interserver_http_port	9009	9009	HTTP port for the communication between ClickHouse servers.
interserver_https_port	9010	9010	HTTPS port for the communication between ClickHouse servers.
http_port	8123	8123	Port for connecting to the ClickHouse server through HTTP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
https_port	8443	8443	Port for connecting to the ClickHouse server through HTTPS.
tcp_port	9000	9000	Port for connecting the client to the ClickHouse server through TCP.
tcp_port_secure	9440	9440	Port for connecting the client to the ClickHouse server through TCP SSL.
lb_tcp_port	21424	21424	TCP communication port number for ClickHouseBalancer.
lb_http_port	21425	21425	HTTP communication port number for ClickHouseBalancer.
lb_https_port	21426	21426	HTTPS communication port number for ClickHouseBalancer.
lb_tcp_secure_port	21428	21428	TCP SSL communication port number for ClickHouseBalancer.

Common Impala Ports

The protocol type of all ports in the table is TCP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
--beeswax_port	21000	21970	Port for impala-shell communication
--hs2_port	21050	21971	Port for Impala application communication
--hs2_http_port	28000	21981	Port used by Impala to provide the HiveServer2 protocol for external systems

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common Doris Ports

The protocol type of the port in the table is TCP and HTTP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
http_port	8030	29980	HTTP port of the FE service
https_port	8050	29991	HTTPS port of the FE service
query_port	9030	29982	Port used by the Doris FE to query connections through the MySQL protocol
rpc_port	9020	29981	Thrift Server port of the FE service
be_port	9060	29984	Thrift Server Port on BE for receiving requests from FE
brpc_port	8060	29987	BRPC port on BE, which is used for communications between BE instances.
heartbeat_service_port	9050	29985	Thrift heartbeat service port on BE, which is used to receive heartbeat messages from FE
webserver_port	8040	29986	HTTP server port on BE
broker_ipc_port	8000	29990	Thrift Server communication port on Broker, which is used to receive requests.
single_replica_load_brpc_port	9070	29988	RPC port used for the communication between the master and slave replicas to import single-replica data
single_replica_load_download_port	8050	29989	Port used by the slave replica to download data files from the master replica through HTTP for single-copy data import

Common Guardian Ports

The protocol type of the port in the table is TCP and HTTP.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Default Open Source Port	Default Custom Port	Port Description
ranger.obs.service.rpc.address.port	26901	26901	The RPC listening port for the Guardian OBS client service.
token.server.http-server.port	26900	26900	HTTP listening port of the Guardian service.
token.server.rpc.address.port	26902	26902	The RPC listening port of the Guardian service.

Common Kudu Ports

The protocol type of the port in the table is TCP and HTTP.

Parameter	Default Open Source Port	Default Custom Port	Port Description
rpc_bind_addresses	7051	7051	Default RPC port of the Kudu Master. All requests from Tablet Server and clients (such as Impala and Spark) interact with the Master through this port.
webserver_port	8051	8051	Web UI port of the Kudu Master. You can access **http://<*master_host*>:8051** using a browser to view the Master status and monitoring information.
rpc_bind_addresses	7050	7050	Default RPC port of the Kudu Tablet Server. This port is used for receiving client requests, coordinating with the Master, and communication between Tablet Servers.
webserver_port	8050	8050	Web UI port of the Kudu Tablet Server. You can access **http://<*tserver_host*>:8050** using a browser to view the Tablet Server status and monitoring information.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

10 Configuring Storage-Compute Decoupling for an MRS Cluster

10.1 MRS Storage-Compute Decoupling Overview

MRS allows you to store data in OBS and use MRS clusters for data computing only. This decoupled storage-compute mode enables you to flexibly scale resources on demand and provides a low-cost solution for massive data analysis.

- MRS 3.3.0- LTS and later versions support only OBS interconnection through the Guardian component, which provides temporary authentication credentials for services such as HDFS, Hive, Spark, Loader, and HetuEngine to access OBS.
  - Configure the Guardian service to disable Ranger OBS path authentication. For details, see **Disabling Ranger OBS Path Authentication for Guardian**.
  - Configure the Guardian service to enable Ranger OBS path authentication and cascading authorization. For details, see **Enabling Ranger OBS Path Authentication for Guardian**.
- For versions earlier than MRS 3.3.0- LTS, OBS interconnection is supported only through the IAM agency mode. For details, see **Interconnecting an MRS Cluster with OBS Using an IAM Agency**.

MRS Storage-Compute Decoupling Solution

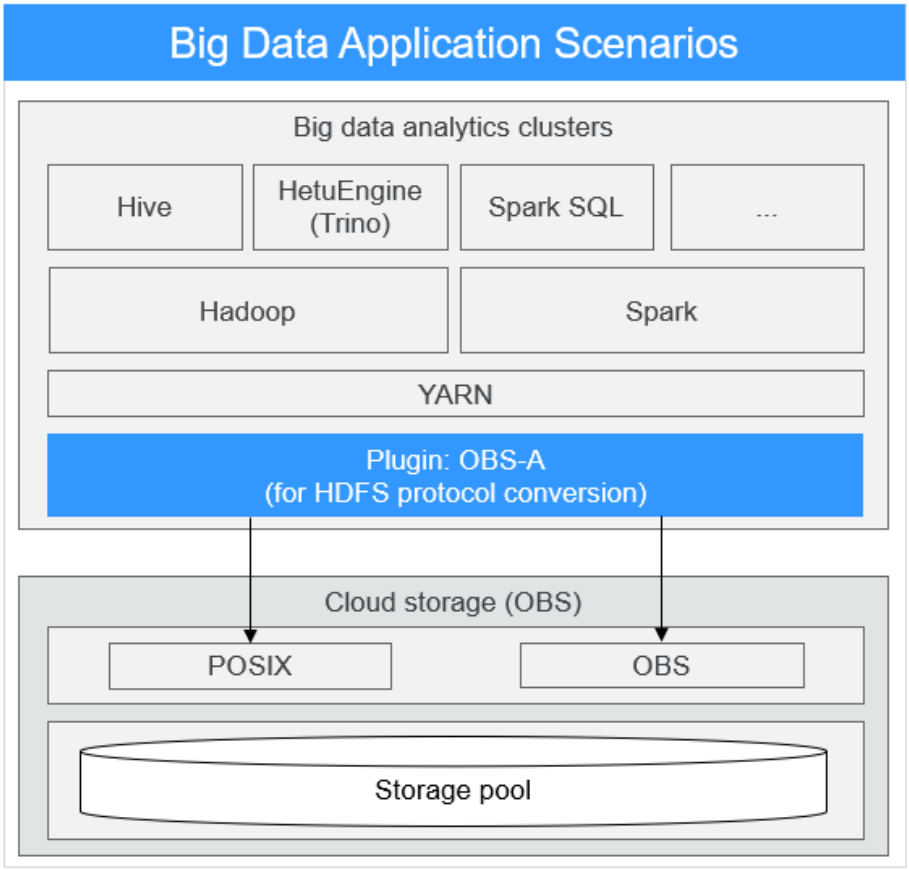
In an MRS cluster, data is stored on local disks or HDFS by default. Data required for real-time computing, stream and batch processing, and interactive analysis is accessed directly from the cluster's local storage and is written back locally after processing. This solution integrates compute and storage.

However, as services grow rapidly and data volumes increase, this solution faces several challenges. These include the high cost of local disk resources, inflexible and unbalanced allocation of compute and storage resources across the cluster, and low disk utilization resulting from Hadoop's default three- replica storage mechanism. MRS provides a storage–compute decoupling solution based on OBS. In scenarios with large data volumes and elastic compute requirements, data can be stored in OBS distributed parallel file system while MRS clusters are used only

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

for computing. This allows resources to be scaled on demand and supports a cost- effective approach to massive data analytics.

Figure 10-1 MRS storage-compute decoupling solution



Based on the large capacity and high bandwidth of OBS, the MRS storage-compute decoupling solution is compatible with multiple computing engines such as Hive and Spark in the Hadoop ecosystem. Compared with traditional on-premises IDC deployment of big data services, the MRS storage-compute decoupling solution features high convergence efficiency, high resource utilization, and low costs.

Notes and Constraints

- In storage-compute decoupling scenarios, make sure to use an OBS parallel file system. Using a regular object bucket can significantly impact the performance of the cluster.
- To delete a component or cluster connected to OBS (including storage-compute decoupling and cold-hot data separation scenarios), you must also delete the service data on OBS.
- After storage-compute decoupling is configured for the MRS cluster, components can access the OBS file system and the HDFS in the cluster. For details, see **Interconnecting an MRS Cluster with OBS Using an IAM Agency** or **Interconnecting an MRS Cluster with OBS Through Guardian**.
- On the Guardian-based storage and compute decoupling management plane, job submission depends on the JobGateway instead of the Executor.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configuring Storage and Compute Decoupling Using Guardian

NOTICE

- If you need to configure permission policies for OBS paths of components using Ranger, that is, enable Ranger OBS path authentication for Guardian, see **Enabling Ranger OBS Path Authentication for Guardian**. For details about OBS permission configuration, see **Granting OBS Permissions**.
- If you only need to configure Guardian to connect to OBS, but do not need to use Ranger to configure permission policies for OBS paths, see **Disabling Ranger OBS Path Authentication for Guardian**.

1. Create an MRS cluster.

   The MRS cluster must contain basic components such as Guardian, Ranger, and Hadoop.

   Only MRS 3.3.0-LTS or later supports interconnection with OBS using Guardian.

2. Create an OBS agency.

   Create an agency with OBS access permissions, which allows Guardian to connect to OBS.

3. Enable the interconnection between Guardian and OBS and configure parameters.

   Modify Guardian service parameters and configure IAM agency authentication information to provide temporary authentication credentials and fine-grained permission control for other components to access OBS.

4. Configure the policy for clearing component data in the recycle bin directory.

   In the storage-compute decoupling scenario, the prevention against accidental deletion is enabled by default for components connected to OBS. When you delete data, the deleted object is moved to the corresponding recycle bin directory. You need to configure a lifecycle rule for the corresponding directory in the OBS file system to prevent the storage space from being used up.

5. Interconnect components with OBS.

6. Components in the MRS cluster can directly access the corresponding path after being granted required permissions for accessing OBS buckets. You can use the component client to directly access resources in the OBS file system in absolute path mode.

Configuring Storage-Compute Decoupling Using IAM Agencies

1. Create an ECS agency with OBS access permissions.

   You need to create an ECS agency with OBS access permissions. The ECS agency is used to automatically obtain temporary AK/SK to access OBS.

2. Create an MRS cluster with decoupled storage and compute.

   - Creating a cluster with decoupled storage and compute

     Create an MRS cluster and bind an ECS agency to it. Add components as required. Components that can be connected to OBS through an IAM

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

agency include Hadoop, Hive, Spark, Presto, Flink, Flume, Hudi and Sqoop.

- Configuring storage-compute decoupling for an existing cluster

  Bind an ECS agency to an existing cluster to implement decoupled storage and compute.

3. Create an OBS file system.

   You need to create an OBS parallel file system for storing cluster data.

4. Configure the policy for clearing component data in the recycle bin directory.

   In the storage-compute decoupling scenario, the prevention against accidental deletion is enabled by default for components connected to OBS. When you delete data, the deleted object is moved to the corresponding recycle bin directory. You need to configure a lifecycle rule for the corresponding directory in the OBS file system to prevent the storage space from being used up.

5. Interconnect components with OBS.
6. Components in the MRS cluster can directly access the corresponding path after being granted required permissions for accessing OBS buckets. You can directly access resources in the OBS file system via the component client using an absolute path.

Granting OBS Permissions

Enabling storage-compute decoupling and Ranger authentication for an MRS cluster that connects to OBS using Guardian allows Ranger administrators to grant cluster users permissions to read and write OBS directories and files.

In addition, with the Guardian-based storage-compute decoupling architecture and the Hive cascading authorization function, you just need to be granted permissions on the service table through Ranger. Then the system will grant you the permissions on the data storage source, for example, the storage directory on OBS, in a fine-grained manner. You do not need to query the storage path of the table and perform secondary authorization.

- Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.
- On Ranger, only users in custom user groups in MRS Manager can be granted OBS permissions. Each user group name must contain 1 to 52 characters and can include only letters, digits, underscores (_), and number signs (#). Otherwise, the policy cannot be added.
- For clusters with Kerberos authentication enabled, permissions need to be granted based on Ranger. For clusters with Kerberos authentication disabled, OBS permissions are granted by default, and no additional configuration is required.

10.2 Interconnecting an MRS Cluster with OBS Through Guardian

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

10.2.1 Disabling Ranger OBS Path Authentication for Guardian

Scenarios

This section describes how to enable storage–compute decoupling for the Guardian component when Kerberos authentication is disabled for the cluster (in normal mode). In this scenario, the Ranger component cannot be used to configure permission policies for OBS paths—that is, the AccessLabel function used by Guardian to interconnect with OBS is disabled, and you do not need to consider whether AccessLabel is enabled on OBS. Once enabled, Guardian can provide temporary authentication credentials for components such as HDFS, Hive, Spark, Loader, and HetuEngine to access OBS in storage-compute decoupling scenarios.

To interconnect Guardian with OBS, do as follows:

1. **Creating an OBS Parallel File System**
2. **Creating a Cloud Service Agency and Binding It to a Cluster**
3. **Granting Guardian the Permission to Access OBS**
4. **Configuring a Recycle Bin Cleanup Policy**

Prerequisites

- Components such as Guardian, Ranger, and Hadoop have been installed in the cluster.
- If components such as Hadoop, HetuEngine, Hive, and Spark have been installed in an environment before installing Guardian, you need to download these component clients again and refresh the default clients for job submission on the MRS console.

Impact on the System

Once you finish the configuration, you will need to either refresh the original client's configuration or reinstall the client.

Creating an OBS Parallel File System

1. Log in to the OBS console.
2. Choose **Parallel File Systems** > **Create Parallel File System**.
3. Enter a file system name, for example, **guardian-obs**.

   Use the enterprise project selected during MRS cluster creation and set other parameters as needed.
4. Click **Create Now**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Creating a Cloud Service Agency and Binding It to a Cluster

NOTE

- MRS presets **MRS_ECS_DEFAULT_AGENCY** in the IAM agency list by default, allowing you to choose this agency when creating a cluster. This agency has **OBS OperateAccess** permission and, for users with fine-grained policies enabled, **CES FullAccess**, **CES Administrator**, and **KMS Administrator** permissions in the region where the cluster is located. Do not modify **MRS_ECS_DEFAULT_AGENCY** on IAM.
- If you want to use the preset agency, skip the step for creating an agency. If you want to use a custom agency, perform the following steps to create an agency. (To create or modify an agency, you must have the **Security Administrator** permission.) If you need to have more fine-grained control over the permissions of a specific path in the OBS file system, you can refer to **Configuring Fine- Grained OBS Access Permissions for MRS Cluster Users** to create a custom role policy.

1. Log in to the management console.
2. In the service list, choose **Management & Governance** > **Identity and Access Management**.
3. Choose **Agencies**. On the displayed page, click **Create Agency**.
4. On the **Create Agency** page, set the following parameters and click **Done**:
   - **Agency Name**: Enter an agency name, for example, **mrs_ecs_obs**.
   - **Agency Type**: Select **Cloud service**.
   - **Cloud Service**: Select **Elastic Cloud Server (ECS) and Bare Metal Server (BMS)**.
   - **Validity Period**: Select **Unlimited**.
5. In the displayed dialog box, click **Authorize**. On the displayed page, click **Create Policy**.

   On the **Create Policy** page, set the following parameters and click **Next**:
   - **Policy Name**: Enter a policy name, for example, **guardian-policy**.
   - **Policy View**: Select **JSON**.
   - **Policy Content**: Configure the parameter as follows:

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
         "obs:bucket:GetBucketLocation",
         "obs:bucket:ListBucketMultipartUploads",
         "obs:object:GetObject",
         "obs:object:ModifyObjectMetaData",
         "obs:object:DeleteObject",
         "obs:object:ListMultipartUploadParts",
         "obs:bucket:HeadBucket",
         "obs:object:AbortMultipartUpload",
         "obs:bucket:ListBucket",
         "obs:object:PutObject",
            "obs:bucket:ListAllMyBuckets"
         ],
         "Resource": [
            "OBS:*:*:bucket:guardian-obs",
            "OBS:*:*:bucket:guardian-obs2",
            "OBS:*:*:object:*"
         ]
      }
   ]
}
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In the preceding configuration, **Resource** indicates all resources that can access the configured parallel file system. *guardian-obs* indicates the name of the OBS parallel file system created in **Creating an OBS Parallel File System**.

6. Click **Next**. On the **Select Policy/Role** page, select the policy created in **5**.
7. Click **Next**, select **All resources**, click **Show More**, select **Global resources**, and click **OK**.
8. In the displayed dialog box, click **OK** to start authorization. After the message "Authorization successful." is displayed, click **Finish**. The agency is successfully created.
9. Log in to the MRS console. In the navigation pane on the left, choose **Active Clusters**.
10. Click the name of the target cluster to access its details page.
11. On the **Dashboard** page, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.
12. On the **Dashboard** tab, click **Select Agency** next to **Agency**. In the displayed dialog box, select the agency you created, for example, **mrs_ecs_obs**, and click **OK**.

Granting Guardian the Permission to Access OBS

1. Log in to MRS Manager by referring to **Accessing MRS Manager**. Choose **Cluster** > **Services** > **Guardian**, click **Configurations** and then **All Configurations**, search for and configure the following parameters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
fs.obs.guardian.accesslabel.enabled	Whether to enable AccessLabel on OBS, which allows Guardian to connect to OBS. • **true**: indicates that after OBS is connected, you can use Ranger to configure the permission policies of OBS paths. Ensure that the AccessLabel function has been enabled for OBS. If the function is not enabled, manually enable it. For details, contact OBS O&M personnel. • **false**: indicates After OBS is interconnected, the permission policies of OBS paths cannot be configured using Ranger. You do not need to check whether the AccessLabel function is enabled for OBS. Set this parameter to **false**.	false
fs.obs.guardian.enabled	Whether to enable Guardian. • **true**: enables the Guardian. • **false**: (default value) disables the Guardian.	true
fs.obs.delegation.token.providers	Delegation token generator. When **fs.obs.guardian.enabled** is set to **true**, **com..mrs.dt.MRSDelegationTokenProvider** and **com..mrs.dt.GuardianDTProvider** must be configured.	**com..mrs.dt.MRSDelegationTokenProvider** and **com..mrs.dt.GuardianDTProvider**

2. Click **Save** to save the service configuration. On the FusionInsight Manager, choose **More** > **Restart Configuration-Expired Instances** and restart all service instances with expired configurations as prompted.
3. To submit jobs on the MRS console, log in to the active OMS node as user **omm** and run the following command to refresh the built-in client configuration:

```
sh /opt/executor/bin/refresh-client-config.sh
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. (Optional) If you need to update the client configurations of other installed cluster components, see **Updating Client Configurations**.

Configuring a Recycle Bin Cleanup Policy

1. Log in to the OBS console.
2. In the navigation pane on the left, choose **Resources** > **Parallel File Systems**. On the displayed page, click the name of the file system created in **Creating an OBS Parallel File System**.
3. Choose **Data Management** > **Lifecycle Rules**. On the displayed page, click **Create** to create a lifecycle rule for the **/user/.Trash** directory.

> **CAUTION**
>
> **After the decoupled storage-compute solution is used, you must configure lifecycle rules for related directories. Otherwise, there is a risk of running out of storage space and incurring additional storage costs.**

Table 10-1 Parameters for creating a lifecycle rule

Parameter	Description	Example Value
Status	Whether to enable the lifecycle rule.	Enabled
Rule Name	Rule name, used to identify different lifecycle configurations.	rule-test
Prefix	Prefix of the objects to which the lifecycle rule applies. Typically, the recycle bin directory of MRS components is prefixed with **/user/.Trash**.	user/.Trash
Transition to Infrequent Access After (Days)	Number of days after the last update of an object that it will be transitioned to infrequent access storage based on the rule. The minimum value is **30**.	30
Transition to Archive After (Days)	Number of days after the last update of an object that it will be transitioned to archive based on the rule. If you are setting both this parameter and **Transition to Infrequent Access After (Days)**, make sure this parameter value is at least 30 days greater than the value of **Transition to Infrequent Access After (Days)**. If you are only setting this parameter, assign any value to it as needed.	31

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Delete Files After (Days)	Number of days after the last update of an object that it will expire and be automatically deleted by OBS based on the rule. The value of this parameter must be greater than the values of the parameters **Transition to Infrequent Access After (Days)** and **Transition to Archive After (Days)**.	32
Delete Fragments Upon Expiration	Number of days of a fragment that it will expire and be automatically deleted by OBS based on the rule.	30

4. Click **OK**.

   To modify the lifecycle configuration, locate the target lifecycle rule, click **Edit** or **Disable** on the right. To enable a lifecycle rule, click **Enable**.

10.2.2 Enabling Ranger OBS Path Authentication for Guardian

Scenarios

This section describes how to configure permission policies for OBS paths through Ranger after the storage-compute decoupling function is enabled and Guardian is connected to OBS. Ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel. Once enabled, Guardian can provide temporary authentication credentials for components such as HDFS, Hive, Spark, Loader, and HetuEngine to access OBS in storage-compute decoupling scenarios.

To interconnect Guardian with OBS, do as follows:

1. **Creating an OBS Parallel File System**
2. **Creating an Agency for a Regular Account**
3. **Creating a Cloud Service Agency and Binding It to a Cluster**
4. **Granting Guardian Permissions to Access OBS**
5. **Enabling Cascading Authorization for Hive Tables**
6. **Configuring a Recycle Bin Cleanup Policy**

Prerequisites

- Components such as Guardian, Ranger, and Hadoop have been installed in the cluster.
- If components such as Hadoop, HetuEngine, Hive, and Spark have been installed in an environment before installing Guardian, you need to download these component clients again and refresh the default clients for job submission on the MRS console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If Kerberos authentication is disabled for the cluster (in normal mode), choose **Cluster** > **Services** > *Service name* on MRS Manager. On the displayed **Dashboard** page, click **More** and select **Enable Ranger** in the upper right corner to enable Ranger authentication for the service.

Impact on the System

- Once you finish the configuration, you will need to either refresh the original client's configuration or reinstall the client.
- To submit a job on console, log in to the active OMS node as user **omm** and run the following command to refresh the cluster's built-in clients.
  ```
  sh /opt/executor/bin/refresh-client-config.sh
  ```

Creating an OBS Parallel File System

1. Log in to the OBS console.
2. Choose **Parallel File Systems** > **Create Parallel File System**.
3. Enter a file system name, for example, **guardian-obs**.

   Use the enterprise project selected during MRS cluster creation and set other parameters as needed.
4. Click **Create Now**.

Creating an Agency for a Regular Account

1. Log in to the management console.
2. In the service list, choose **Management & Governance** > **Identity and Access Management**.
3. Choose **Agencies**. On the displayed page, click **Create Agency**.
4. On the **Create Agency** page, set the following parameters and click **Done**:
   - **Agency Name**: Enter an agency name, for example, **agency-MRS-to-OBS**.
   - **Agency Type**: Select **Account**.
   - **Delegated Account**: Enter your cloud account that you signed up for with your mobile number. It cannot be a federated user or an IAM user created using your cloud account.
   - **Validity Period**: Select **Unlimited**.
5. In the displayed dialog box, click **Authorize**. On the displayed page, click **Create Policy**.

   On the **Create Policy** page, set the following parameters and click **Next**:
   - **Policy Name**: Enter a policy name, for example, **guardian-policy**.
   - **Policy View**: Select **JSON**.
   - **Policy Content**: Configure the parameter as follows:
     ```
     {
        "Version": "1.1",
        "Statement": [
           {
              "Effect": "Allow",
              "Action": [
              "obs:bucket:GetBucketLocation",
              "obs:bucket:ListBucketMultipartUploads",
              "obs:object:GetObject",
     ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
            "obs:object:ModifyObjectMetaData",
            "obs:object:DeleteObject",
            "obs:object:ListMultipartUploadParts",
            "obs:bucket:HeadBucket",
            "obs:object:AbortMultipartUpload",
            "obs:bucket:ListBucket",
            "obs:object:PutObject",
            "obs:object:GetAccessLabel",
            "obs:object:DeleteAccessLabel",
            "obs:object:PutAccessLabel",
                "obs:bucket:ListAllMyBuckets"
            ],
            "Resource": [
                "OBS:*:*:bucket:guardian-obs",
                "OBS:*:*:object:*"
            ]
        }
    ]
}
```

In the preceding configuration, **Resource** indicates all resources that can access the configured parallel file system. *guardian-obs* indicates the name of the OBS parallel file system created in **Creating an OBS Parallel File System**.

6. Click **Next**. On the **Select Policy/Role** page, select the policy created in **5**.
7. Click **Next**, select **All resources**, click **Show More**, select **Global resources**, and click **OK**.
8. View and record the agency ID.

Creating a Cloud Service Agency and Binding It to a Cluster

1. Log in to the management console.
2. In the service list, choose **Management & Governance** > **Identity and Access Management**.
3. Choose **Agencies**. On the displayed page, click **Create Agency**.
4. Set **Agency Name**. For example, enter **mrs_ecs_obs**.
5. Set **Agency Type** to **Cloud service** and select **Elastic Cloud Server (ECS) and Bare Metal Server (BMS)** to authorize ECS or BMS to access OBS.
6. Set **Validity Period** to **Unlimited** and click **Done**.
7. In the displayed dialog box, click **Authorize**. On the displayed page, click **Create Policy**.

   On the **Create Policy** page, set the following parameters and click **Next**:

   - **Policy Name**: Enter a policy name, for example, **guardian-assume-policy**.
   - **Policy View**: Select **JSON**.
   - **Policy Content**: Enter the following content. *{Agency ID}* indicates the ID recorded in **8**.

```
{
    "Version": "1.1",
    "Statement": [
        {
            "Action": [
                "iam:agencies:assume"
            ],
            "Resource": {
                "uri": [
                    "/iam/agencies/{Agency ID}"
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
            ]
        },
        "Effect": "Allow"
    }
  ]
}
```

8. Click **Next**. On the **Select Policy/Role** page, select the policy created in 7.
9. Click **Next**, click **Show More**, select **Global services**, and click **OK**.
10. In the displayed dialog box, click **OK** to start authorization. Click **Finish** after the message "Authorization successful." is displayed.
11. Log in to the MRS console. In the navigation pane on the left, choose **Active Clusters**.
12. Click the name of the target cluster to access its details page.
13. On the **Dashboard** tab, click **Synchronize** next to **IAM User Sync** to synchronize IAM users.
14. On the **Dashboard** tab, click **Manage Agency** next to **Agency**. In the displayed dialog box, select the agency you created, for example, **mrs_ecs_obs**, and click **OK**.

Granting Guardian Permissions to Access OBS

1. Log in to MRS Manager by referring to **Accessing MRS Manager**. Choose **Cluster** > **Services** > **Guardian**, click **Configurations** and then **All Configurations**, search for and configure the following parameters.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
fs.obs.guardian.accesslabel.enabled	Whether to enable AccessLabel on OBS, which allows Guardian to connect to OBS. • **true**: indicates that after OBS is connected, you can use Ranger to configure the permission policies of OBS paths. Ensure that the AccessLabel function has been enabled for OBS. If the function is not enabled, manually enable it. For details, contact OBS O&M personnel. • **false**: indicates After OBS is interconnected, the permission policies of OBS paths cannot be configured using Ranger. You do not need to check whether the AccessLabel function is enabled for OBS. Set this parameter to **true**.	true
fs.obs.guardian.enabled	Whether to enable Guardian. • **true**: enables the Guardian. • **false**: (default value) disables the Guardian.	true
fs.obs.delegation.token.providers	Delegation token generator. When **fs.obs.guardian.enabled** is set to **true**, **com..mrs.dt.MRSDelegationTokenProvider** and **com..mrs.dt.GuardianDTProvider** must be configured.	**com..mrs.dt.MRSDelegationTokenProvider** and **com..mrs.dt.GuardianDTProvider**
token.server.access.label.agency.name	Name of the specified IAM agency, which is the one created in **Creating an Agency for a Regular Account**.	agency-MRS-to-OBS

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Save the service configuration, choose **More** > **Restart Configuration-Expired Instances** on the FusionInsight Manager home page, and restart all service instances whose configurations have expired as prompted.
3. To submit jobs on the MRS console, log in to the active OMS node as user **omm** and run the following command to refresh the built-in client configuration:
```
sh /opt/executor/bin/refresh-client-config.sh
```
4. (Optional) If you need to update the client configurations of other installed cluster components, see **Updating Client Configurations**.

Enabling Cascading Authorization for Hive Tables

In versions earlier than MRS 3.5.0, only clusters with Kerberos authentication enabled support OBS cascading authorization.

1. Log in to FusionInsight Manager by referring to **Accessing MRS Manager**. Choose **Cluster** > **Services** > **Ranger** > **Configurations**.
2. Search for **ranger.ext.authorization.cascade.enable** in the search box and set its value to **true** to enable the cascading authorization.
3. Click **Save**.
4. Click **Instance** and select all RangerAdmin instances. Click **More** and select **Restart Instance**. Enter the password and click **OK** to restart all RangerAdmin instances.

Configuring a Recycle Bin Cleanup Policy

1. Log in to the OBS console.
2. In the navigation pane on the left, choose **Resources** > **Parallel File Systems**. On the displayed page, click the name of the file system created in **Creating an OBS Parallel File System**.
3. Choose **Data Management** > **Lifecycle Rules**. On the displayed page, click **Create** to create a lifecycle rule for the **/user/.Trash** directory.

CAUTION

After the decoupled storage-compute solution is used, you must configure lifecycle rules for related directories. Otherwise, there is a risk of running out of storage space and incurring additional storage costs.

Table 10-2 Parameters for creating a lifecycle rule

Parameter	Description	Example Value
Status	Whether to enable the lifecycle rule.	Enabled
Rule Name	Enter a rule name, which is used to identify different lifecycle configurations.	rule-test

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Prefix	Prefix of the objects to which the lifecycle rule applies. Typically, the prefix of the recycle bin directory of MRS components is **/user/.Trash**.	user/.Trash
Transition to Infrequent Access After (Days)	Number of days after the last update of an object that it will be transitioned to infrequent access storage based on the rule. The minimum value is **30**.	30 days
Transition to Archive After (Days)	Number of days after the last update of an object that it will be transitioned to archive based on the rule. If you are setting both this parameter and **Transition to Infrequent Access After (Days)**, make sure this parameter value is at least 30 days greater than the value of **Transition to Infrequent Access After (Days)**. If you are only setting this parameter, assign any value to it as needed.	31 days
Delete Files After (Days)	Number of days after the last update of an object that it will expire and be automatically deleted by OBS based on the rule. The value of this parameter must be greater than the values of the two transition parameters.	32 days
Delete Fragments Upon Expiration	Number of days of a fragment that it will expire and be automatically deleted by OBS based on the rule.	30 days

4. Click **OK**.

   To modify, disable, or enable a lifecycle rule, locate the rule and click **Edit**, **Disable**, or **Enable** in the **Operation** column, respectively.

10.2.3 Example for Interconnecting an MRS Cluster with OBS

10.2.3.1 Accessing OBS Using Flink Through Guardian

After Guardian is interconnected with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you can access the OBS parallel file system and run jobs in a Flink job. This section describes how to submit a Flink WordCount job.

Prerequisites

- Before interconnecting Flink with OBS, ensure that YARN is connected to OBS as Flink jobs run on YARN.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The provided OBS parallel file system name/file name must include the full directory hierarchy.
- If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Ranger Permission Configuration**.

Interconnecting Flink with OBS

Step 1 Log in to the Flink client installation node as the client installation user.

Step 2 Run the following command to initialize environment variables:

```
source Client installation directory/bigdata_env
```

Step 3 Configure the Flink client.

Step 4 Start the session.

- Normal cluster (Kerberos authentication disabled)

  ```
  yarn-session.sh -nm "session-name" -d
  ```

- Security cluster (Kerberos authentication enabled)
  - If the paths of the **flink.keystore** and **flink.truststore** files are relative ones:

    Run the following command in the directory at the same level as the **ssl** directory to start the session. **ssl/** is a relative path.

    Switch to the **conf** directory.

    ```
    cd Client installation directory/Flink/flink/conf
    ```

    Start a session.

    ```
    yarn-session.sh -t ssl/ -nm "session-name" -d
    ```

    The session starts successfully.

    ```
    ...
    Cluster started: Yarn cluster with application id application_1624937999496_0017
    JobManager Web Interface: http://192.168.1.150:32261
    ```

  - If the paths of the **flink.keystore** and **flink.truststore** files are absolute ones, run the following commands to start a session:

    Switch to the **conf** directory.

    ```
    cd Client installation directory/Flink/flink/conf
    ```

    Start a session.

    ```
    yarn-session.sh -nm "session-name" -d
    ```

Step 5 If Kerberos authentication is enabled for the cluster, run the following command to authenticate the user. Skip this step if Kerberos authentication is disabled.

```
kinit Username
```

Step 6 Add the target OBS file system to the Flink CLI and execute the analysis program.

1. Create the **test** file in the **/tmp/test** directory.

   ```
   echo -e 'test' >/tmp/test
   ```

2. Create a directory for storing the **test** file in the OBS parallel file system.

   ```
   hdfs dfs -mkdir -p obs://Parallel file system name/tmp/flinkjob
   ```

3. Upload the **test** file to the OBS parallel file system.

   ```
   hdfs dfs -put /tmp/test/ obs://Parallel file system name/tmp/flinkjob/
   ```

4. Run **WordCount.jar**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
flink run Client installation directory/Flink/flink/examples/batch/WordCount.jar -input obs://Parallel file system name/tmp/flinkjob/test -output obs://Parallel file system name/tmp/flinkjob/output
```

Step 7 Log in to the OBS console and check the execution result of **WordCount.jar** in the output path specified in **Step 6.4**.

----End

Ranger Permission Configuration

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group**.

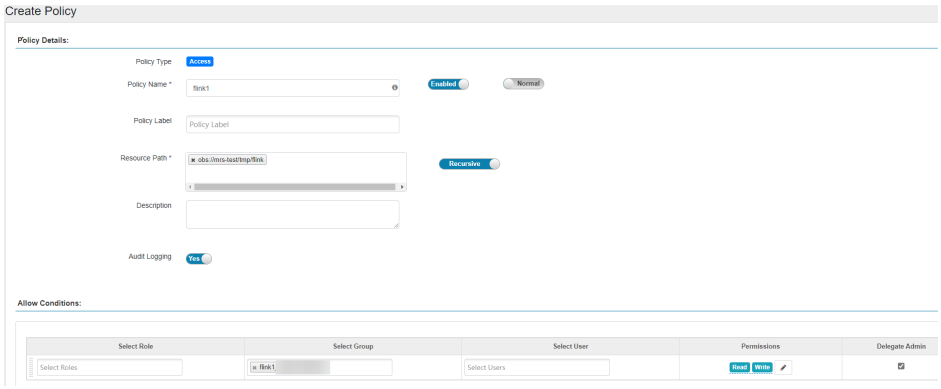
Step 3 Create a user group without a role, for example, **obs_flink**, and bind the user group to the corresponding user.

Step 4 Log in to the Ranger management page as user **rangeradmin**.

Step 5 On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

Step 6 Click **Add New Policy** to add the **Read** and **Write** permissions on OBS paths to the user group created in **3**. If there are no OBS paths, create one in advance. The OBS path cannot contain wildcard characters (*).

Figure 10-2 Granting the Flink user group permissions for reading and writing OBS paths



Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

----End

10.2.3.2 Accessing OBS Using HDFS Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Authentication for Guardian, you can view and create OBS file directories on the HDFS client.

Prerequisites

If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Configuring Ranger Permissions**.

Interconnecting HDFS with OBS

Step 1 Log in to the node on which the HDFS client is installed as a client installation user.

Step 2 Run the following command to switch to the client installation directory:

```
cd Client installation directory
```

Step 3 Run the following command to configure environment variables:

```
source bigdata_env
```

Step 4 If Kerberos authentication is enabled for the cluster, run the following command to perform user authentication. The user must have the read and write permissions on the OBS directory. User authentication is not required for clusters with Kerberos authentication disabled.

```
kinit User performing HDFS operations
```

Step 5 Access the OBS file system using the HDFS CLI.

The following commands are examples.

- Access the OBS file system.
  ```
  hdfs dfs -ls obs://OBS parallel file system name/path
  ```
- Create a directory in the OBS file system.
  ```
  hdfs dfs -mkdir obs://OBS parallel file system name/hadoop
  ```
- Upload the **/opt/test.txt** file on the client node to the **obs://***OBS parallel file system name***/hadoop** directory.
  ```
  hdfs dfs -put /opt/test.txt obs://OBS parallel file system name/hadoop
  ```
  View the uploaded file in the corresponding OBS file system path.

----End

Changing the Log Level of OBS Client

If a large number of logs are printed in the OBS file system, the read and write performance may be affected. You can adjust the log level of the OBS client as follows:

1. Go to the **hadoop** directory.
   ```
   cd Client installation directory/HDFS/hadoop/etc/hadoop
   ```
2. Edit the file **log4j.properties**.
   ```
   vi log4j.properties
   ```
   Add the following OBS log level configuration to the file and save it.
   ```
   log4j.logger.org.apache.hadoop.fs.obs=WARN
   log4j.logger.com.obs=WARN
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Run the following command:
   ```
   tail -4 log4j.properties
   ```
   If the command output shown in **Figure 10-3** is displayed, the log level has been changed.

   Figure 10-3 Adding an OBS log level

   ```
   [root@node-master1AuKK hadoop]# tail -4 log4j.properties
   log4j.logger.org.apache.commons.beanutils=WARN

   log4j.logger.org.apache.hadoop.fs.obs=WARN
   log4j.logger.com.obs=WARN
   [root@node-master1AuKK hadoop]#
   ```

Configuring Ranger Permissions

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group** to create a user group without any roles, for example, **obs_hadoop**.

Step 3 Back to FusionInsight Manager and choose **System** > **Permission** > **User**. On the displayed page, click **Create User** to create a user that is associated only with the **obs_hadoop** user group, for example, **hadoopuser**.

Step 4 Log in to the Ranger management page as the **rangeradmin** user.

Step 5 On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

Step 6 Click **Add New Policy** and add the **Read** and **Write** permissions on the desired OBS paths to the created user group.

The following figure shows the configurations needed for adding the **Read** and **Write** permissions on **obs://***OBS parallel file system name***/hadoop** to user group **obs_hadoop**.

Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

----End

10.2.3.3 Accessing OBS Using HetuEngine Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you can create a table with a specified location as the OBS file system path on the HetuEngine client. Additionally, you can interconnect Guardian with OBS using Hive MetaStore.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Configuring Ranger Permissions**.

Interconnecting HetuEngine with OBS

In an MRS cluster, **Location** can be set to an OBS file system path during HetuEngine table creation and HetuEngine can connect to OBS through Hive Metastore.

- Setting **Location** to the OBS file system path when creating a table
  a. If a HetuEngine compute instance is running, restart it.

     Log in to FusionInsight Manager as a user who has permission to access the HetuEngine web UI. Choose **Cluster** > **Services** > **HetuEngine**. In the **Basic Information** area in the Dashboard tab, click the link next to **HSConsole WebUI**. On the displayed HSConsole page, click **Compute Instance**. In the instance list, click **Restart** in the **Operation** column and operate as prompted.

  b. Log in to the node where the HetuEngine service client is located as the client installation user and run the following command:
     ```
     source Client installation directory/bigdata_env
     ```

  c. Log in to the HetuEngine client.

     - If the cluster has enabled Kerberos authentication (security mode), run the following command to complete user authentication and log in to the HetuEngine client:

       Authenticate the user.

       ```
       kinit User performing HetuEngine operations
       ```

       Log in to the HetuEngine client.

       ```
       hetu-cli --catalog hive --tenant default --schema default
       ```

     - Kerberos authentication is not enabled for the cluster (normal mode): Run the following command to log in to the HetuEngine client:
       ```
       hetu-cli --catalog hive --tenant default --schema default --user User performing HetuEngine operations
       ```

  d. Set **Location** to the OBS file system path when creating a table.
     ```
     create table test(name string) with (location = 'obs://OBS parallel file system name/user/hive/warehouse/test');
     ```

- Interconnecting with OBS through Hive Metastore
  a. Complete the configurations by referring to **Interconnecting Hive with OBS using MetaStore**.

  b. Log in to FusionInsight Manager, choose **Cluster** > **Services** > **HetuEngine**. On the displayed page, choose **More** > **Synchronize Configuration**. After the synchronization is complete, choose **More** > **Synchronize Configuration** again and then restart the HetuEngine service as prompted.

     If a HetuEngine compute instance is running, stop it before restarting the service. After the service is restarted, start this compute instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

c. No location needs to be specified when you log in to the HetuEngine client to create a schema or table. The schema or table is stored on OBS by default.

Configuring Ranger Permissions

For HetuEngine clusters with Kerberos authentication enabled (security mode), the methods to grant Ranger permission are the same for both storage-compute decoupled architecture and storage-compute coupled architecture.

Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

10.2.3.4 Accessing OBS Using Hive Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you can create tables stored in the OBS parallel file system on the Hive client.

Prerequisites

If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Configuring Ranger Permissions**.

Interconnecting Hive with OBS

MRS clusters allow Hive to connect to OBS through Metastore.

Interconnecting Hive with OBS through Metastore

Step 1 You have configured storage and compute decoupling by referring to **Enabling Ranger OBS Path Authentication for Guardian**.

Step 2 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 3 Choose **Cluster** > **Services** > **Hive** and click **Configurations**.

Step 4 Search for **hive.metastore.warehouse.dir** and set its value to an OBS path, for example, **obs://hivetest/user/hive/warehouse/**. **hivetest** indicates the OBS file system name.

Step 5 Click **Save** to save the configuration, choose **Cluster** > **Services**, and restart the Hive service in the service list.

Step 6 Update the client configuration file.

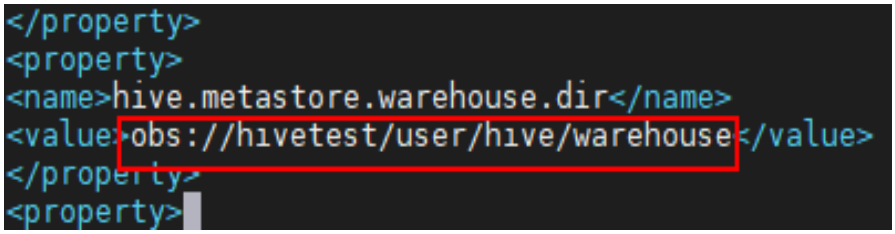
1. Log in to the node where the Hive client is located and run the following command to modify **hivemetastore-site.xml** in the Hive client configuration file directory:

```
vi Client installation directory/Hive/config/hivemetastore-site.xml
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Change the value of **hive.metastore.warehouse.dir** to the corresponding OBS path, for example, **obs://hivetest/user/hive/warehouse/**.

   Figure 10-4 Configuring the OBS path

3. Change the value of **hive.metastore.warehouse.dir** of **hivemetastore-site.xml** in the HCatalog client configuration file directory to the corresponding OBS path, for example, **obs://hivetest/user/hive/warehouse/**.

   ```
   vi Client installation directory/Hive/HCatalog/conf/hivemetastore-site.xml
   ```

Step 7 Go to the Hive Beeline CLI, create a database, and ensure that the location is an OBS path.

1. Go to the client installation directory.

   ```
   cd Client installation directory
   ```

2. Load the environment variables.

   ```
   source bigdata_env
   ```

3. Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

   ```
   kinit User performing Hive operations
   ```

4. Log in to the Hive client.

   ```
   beeline
   ```

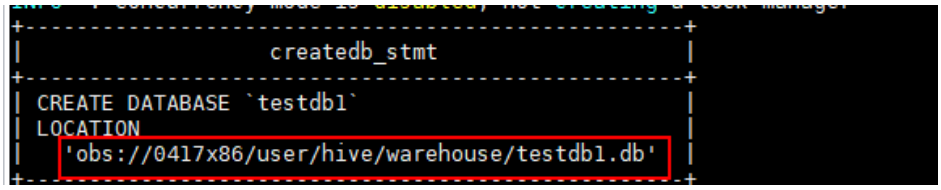
5. Create a database.

   ```
   create database testdb1;
   ```

6. Check the location of the database.

   ```
   show create database testdb1;
   ```

   Figure 10-5 Viewing the location of the newly created Hive database

----End

Configuring Ranger Permissions

- Granting the read and write permissions on OBS paths to the **hive** user group

  a. Log in to the Ranger web UI as the Ranger administrator. On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area, and assign the **Read** and **Write** permissions on the OBS storage path to the **hive** user group. If this operation is successful, all users in the **hive** group can access the Hive data warehouse path.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For example, assign the **Read** and **Write** permissions on the **obs://hivetest/user/hive/warehouse/** directory to the **hive** user group.

b. Choose **Settings** > **Roles**, click **Add New Role**, and create a role whose **Role Name** is **hive**.

- Granting the read and write permissions on OBS paths to a custom user group

  a. Log in to FusionInsight Manager and choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group**.

  b. Create a user group without a role, for example, **hiveobs1**, and bind the user group to the corresponding user.

  c. Log in to the Ranger management page as user **rangeradmin**.

  d. On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

  e. Grant the **Read** and **Write** permissions on the OBS storage path to the **hiveobs1** user group. In this case, all users bound to the **hiveobs1** user group can access the Hive data warehouse path.

- Creating a database, table, or partition in a custom location and granting read and write permissions on OBS paths

  a. Log in to the Ranger web UI as the Ranger administrator.

  b. On the home page, click the component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area, and assign the **Read** and **Write** permissions on OBS storage paths to the user group of the corresponding user.

  For example, assign the **Read** and **Write** permissions on the **obs://obs-test/test/** directory to the **hgroup1** user group, as shown in the following figure.

  c. On the home page, click the component plug-in name **Hive** in the **HADOOP SQL** area, and add a URL policy that grants the **Read** and **Write** permissions on the OBS path to the user group of the corresponding user.

  For example, create the **hive_url_policy URL** policy for the **hgroup1** user group and assign the **Read** and **Write** permissions on the **obs://obs-test/test/** directory to the user group.

  d. Log in to the beeline client and set **Location** to the OBS file system path when creating a table.

  Go to the client installation directory.

  ```
  cd Client installation directory
  ```

  Load the environment variables.

  ```
  source bigdata_env
  ```

  Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

  ```
  kinit User performing Hive operations
  ```

  Log in to the Hive client.

  ```
  beeline
  ```

  For example, to create a table named **test** whose **Location** is **obs://obs-test/test/***Database name*/*Table name*, run the following command:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
create external table test(name string) location " obs://obs-test/test/ Database name/Table name";
```

NOTE

- To authorize a view chart, you need to grant the view chart permission and the physical table path permission corresponding to the view chart.
- Cascading authorization can be performed only on databases and tables, and cannot be on partitions. If a partition path is not in the table path, you need to manually authorize the partition path.
- Cascading authorization for **Deny Conditions** in the Hive Ranger policy is not supported. That is, the Deny Conditions permission only restricts the table permission and cannot generate the permission of the HDFS/OBS storage source.
- The permission of the HDFS Ranger policy is prior to that of the HDFS/OBS storage source generated by cascading authorization. If the HDFS Ranger permission has been set for the HDFS storage source of the table, the cascading permission does not take effect.
- **alter** operations cannot be performed on tables whose storage source is OBS after cascading authorization. To perform the **alter** operation, you need to grant the **Read** and **Write** permissions of the parent directory of the OBS table path to the corresponding user group.
- Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

10.2.3.5 Accessing OBS Using Hudi Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you can create a Hudi COW table in spark-shell and store it to OBS.

Prerequisites

If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Configuring Ranger Permissions**.

Interconnecting Hudi with OBS

Step 1 Log in to the node where the client is installed as the client installation user.

Step 2 Run the following commands to configure environment variables:

Load the environment variables.

```
source Client installation directory/bigdata_env
```

Load the component environment variables.

```
source Client installation directory/Hudi/component_env
```

Step 3 Modify the configuration file.

```
vim Client installation directory/Hudi/hudi/conf/hdfs-site.xml
```

Modify the following content, where the **dfs.namenode.acls.enabled** parameter specifies whether to enable the HDFS ACL function.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
<property>
<name>dfs.namenode.acls.enabled</name>
<value>false</value>
</property>
```

Step 4 If Kerberos authentication has been enabled (security mode) for the cluster, run the following command to perform authentication as a user who has the Read and Write permissions on the corresponding OBS path. If Kerberos authentication has not been enabled (normal mode) for the cluster, you do not need to run this command.

```
kinit Username
```

Step 5 Start **spark-shell** and run the following commands to create a COW table and store it to OBS:

```
spark-shell --master yarn
import org.apache.hudi.QuickstartUtils._
import scala.collection.JavaConversions._
import org.apache.spark.sql.SaveMode._
import org.apache.hudi.DataSourceReadOptions._
import org.apache.hudi.DataSourceWriteOptions._
import org.apache.hudi.config.HoodieWriteConfig._
val tableName = "hudi_cow_table"
val basePath = "obs://testhudi/cow_table/"
val dataGen = new DataGenerator
val inserts = convertToStringList(dataGen.generateInserts(10))
val df = spark.read.json(spark.sparkContext.parallelize(inserts, 2))
df.write.format("org.apache.hudi").
options(getQuickstartWriteConfigs).
option(PRECOMBINE_FIELD_OPT_KEY, "ts").
option(RECORDKEY_FIELD_OPT_KEY, "uuid").
option(PARTITIONPATH_FIELD_OPT_KEY, "partitionpath").
option(TABLE_NAME, tableName).
mode(Overwrite).
save(basePath);
```

In the preceding command, **obs://testhudi/cow_table/** indicates the OBS path, and **testhudi** indicates the OBS parallel system file name. Change them as required.

Step 6 Use DataSource to check whether the table is created and whether the data is normal.

```
val roViewDF = spark.
read.
format("org.apache.hudi").
load(basePath + "/*/*/*/*")
roViewDF.createOrReplaceTempView("hudi_ro_table")
spark.sql("select * from  hudi_ro_table").show()
```

Step 7 Exit the **spark-shell** CLI.

```
:q
```

----End

Configuring Ranger Permissions

Step 1 Log in to FusionInsight Manager and choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Create a user group without a role, for example, **obs_hudi**, and bind the user group to the corresponding user.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Log in to the Ranger management page as the **rangeradmin** user.

Step 4 On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

Step 5 Click **Add New Policy** to add the **Read** and **Write** permissions on OBS paths to the user group created in **Step 2**. If there are no OBS paths, create one in advance (wildcard character * is not allowed).

Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

----End

10.2.3.6 Accessing OBS Using MapReduce Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you need to add custom configurations for the MapReduce component.

Interconnecting MapReduce with OBS

Step 1 Log in to the MRS management console. In the navigation pane on the left, choose **Clusters** > **Active Clusters**. On the displayed page, click the name of the cluster you created to access its details page.

Step 2 Choose **Components** > **MapReduce**.On the displayed page, click **Configurations** then **All Configurations**. In the navigation pane on the left, choose **MapReduce** > **Customization**. In the custom configuration items, add the configuration item **mapreduce.jobhistory.always-scan-user-dir** to **core-site.xml** file.

This parameter specifies whether the JobHistory service always scans log files in the user directory when scanning job logs. The value can be **true** or **false**.

- **true**: The JobHistory service forcibly scans the specified directory even if the directory has been scanned. This ensures that all related log files are correctly collected and processed each time a job is executed.
- **false** (default value): The JobHistory service does not scan the user directories that have been scanned. This reduces unnecessary disk I/O operations.

Set this parameter to **true**.

Step 3 Click **Save Configuration**.

Step 4 Click the **Service Status** tab and choose **More** > **Restart Service** to restart the MapReduce service.

Step 5 If the OBS job is submitted and executed, MapReduce is successfully connected to OBS. For example, create a Hive table and set its **Location** to an OBS path.

1. Log in to the node where the Hive client is installed and run the following commands. The authenticated user must have the read and write permissions to create Hive tables and access OBS paths.

   Go to the client installation directory.

   ```
   cd Client installation directory
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Load the environment variables.

```
source bigdata_env
```

Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

```
kinit Service user
```

2. Log in to the Hive client.

```
beeline
```

3. Create a Hive table and insert data into the table.

   a. Create a table.

   ```
   create table test(name string) location "obs://OBS parallel file system name/user/hive/warehouse/Database name/Table name";
   ```

   b. Check the **Location** of the table.

   ```
   desc formatted test;
   ```

   It is found that the **Location** of the table is the OBS path.

   Figure 10-6 Checking the table location

   ```
   +-------------------------------+-----------------------------------------------------+-----------------------+
   |           col_name            |                      data_type                      |        comment        |
   +-------------------------------+-----------------------------------------------------+-----------------------+
   | # col_name                    | data_type                                           | comment               |
   | name                          | string                                              |                       |
   |                               | NULL                                                | NULL                  |
   | # Detailed Table Information  | NULL                                                | NULL                  |
   | Database:                     | default                                             | NULL                  |
   | OwnerType:                    | USER                                                | NULL                  |
   | Owner:                        | root                                                | NULL                  |
   | CreateTime:                   | Fri May 16 18:04:09 CST 2025                        | NULL                  |
   | LastAccessTime:               | UNKNOWN                                             | NULL                  |
   | Retention:                    | 0                                                   | NULL                  |
   | Location:                     | obs://                                              | NULL                  |
   | Table Type:                   | MANAGED_TABLE                                       | NULL                  |
   | Table Parameters:             | NULL                                                | NULL                  |
   |                               | bucketing_version                                   | 2                     |
   |                               | transient_lastDdlTime                               | 1747389049            |
   |                               | NULL                                                | NULL                  |
   | # Storage Information         | NULL                                                | NULL                  |
   | SerDe Library:                | org.apache.hadoop.hive.serde2.columnar.ColumnarSerDe | NULL                  |
   | InputFormat:                  | org.apache.hadoop.hive.ql.io.RCFileInputFormat      | NULL                  |
   | OutputFormat:                 | org.apache.hadoop.hive.ql.io.RCFileOutputFormat     | NULL                  |
   | Compressed:                   | No                                                  | NULL                  |
   | Num Buckets:                  | -1                                                  | NULL                  |
   | Bucket Columns:               | []                                                  | NULL                  |
   | Sort Columns:                 | []                                                  | NULL                  |
   | Storage Desc Params:          | NULL                                                | NULL                  |
   |                               | serialization.format                                | 1                     |
   +-------------------------------+-----------------------------------------------------+-----------------------+
   ```

   c. Insert data.

   ```
   insert into table test values("A");
   ```

   d. Query table data.

   ```
   select * from test;
   ```

4. Log in to FusionInsight Manager, choose **Cluster** > **Services** > **Yarn**, and click the hyperlink next to **ResourceManager WebUI** to go to the YARN web UI. On the **All Applications** page, click the application ID of the target Hive task to view the job running information. If **Application Type** is **MAPREDUCE**, the job submitted to OBS is a MapReduce job.
5. Click **Logs** in the **Logs** column at the bottom of the page to view job logs.

----End

10.2.3.7 Accessing OBS Using Spark Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you can create a table whose location is an OBS path on the Spark client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Prerequisites

If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Configuring Ranger Permissions**.

Interconnecting Spark with OBS

In an MRS cluster, **Location** can be set to an OBS file system path during Spark table creation and Spark can connect to OBS through Hive Metastore.

- Setting the location to an OBS path during table creation:
  a. Log in to the node where the client is installed as the client installation user and access the **spark-sql** client.

     Go to the client installation directory.

     ```
     cd Client installation directory
     ```

     Load the environment variables.

     ```
     source bigdata_env
     ```

     Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

     ```
     kinit User performing operations on the component
     ```

     Log in to the **spark-sql** client.

     ```
     spark-sql --master yarn
     ```

  b. Set **Location** to the OBS file system path when creating a table.

     For example, to create a table named **test** whose **Location** is **obs://obs-test/test/***Database name***/***Table name*, run the following command:

     ```
     create external table testspark(name string) location " obs://obs-test/test/ Database name/ Table name";
     ```

- Interconnecting Spark with OBS through Hive Metastore:
  a. Complete the configurations by referring to **Interconnecting Hive with OBS using MetaStore**.
  b. Log in to FusionInsight Manager, choose **Cluster** > **Services** > **Spark** and choose **Configurations** > **All Configurations**.
  c. In the navigation pane on the left, choose **SparkResource** > **Customization**. In the custom configuration items, add **spark.sql.warehouse.location.first** to the **custom** parameter and set its value to **true**.
  d. In the navigation pane on the left, choose **JDBCServer** > **Customization**. In the custom configuration items, add **spark.sql.warehouse.location.first** to the **custom** parameter and set its value to **true**.
  e. Click **Save** to save the configuration. Click the **Dashboard** tab choose **More** > **Restart Service**, enter the password, click **OK**, and click **OK** again to restart Spark.
  f. After Spark is restarted, choose **More** > **Download Client** to download and install the Spark client again. Then, go to **g**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If you do not download and install the client again, you can perform the following steps to update the Spark client configuration file (for example, the client directory is **/opt/client**):

i. Log in to the node where the Spark client is deployed as user **root** and switch to the client installation directory.
```
cd /opt/client
```

ii. Run the following command to modify **hive-site.xml** in the configuration file directory of the Spark client:
```
vi Spark/spark/conf/hive-site.xml
```
Change the value of **hive.metastore.warehouse.dir** to the corresponding OBS path, for example, **obs://hivetest/user/hive/warehouse/**.
```
<property>
<name>hive.metastore.warehouse.dir</name>
<value>obs://hivetest/user/hive/warehouse/<value>
</property>
```

iii. Run the following command to modify the **spark-defaults.conf** file in the configuration file directory of the Spark client and set **spark.sql.warehouse.location.first** to **true**:
```
vi Spark/spark/conf/spark-defaults.conf
```

g. For a cluster with Kerberos authentication enabled, configure the OBS directory permission for the user who performs operations on the component. For details, see **Configuring Ranger Permissions**.

h. Go to the SparkSQL CLI and spark-beeline, create a table, and check whether the location of the table is an OBS path.

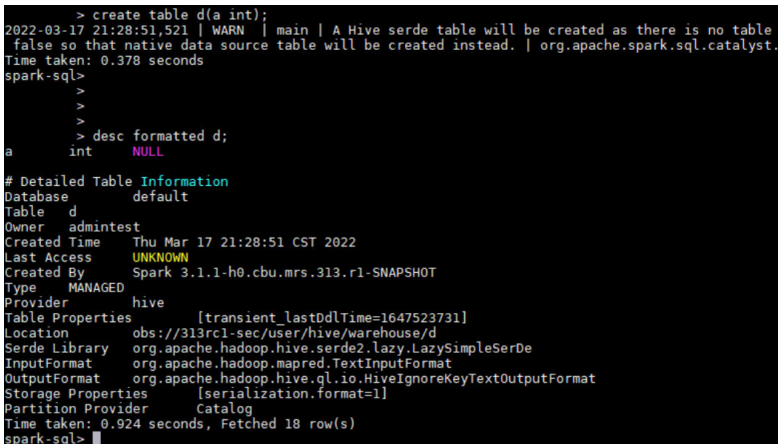
Load the environment variables.
```
source bigdata_env
```

Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.
```
kinit Service user
```

- Go to the SparkSQL CLI.
```
spark-sql
```
Create a table.
```
create table d(a int);
```
Check the table location.
```
desc formatted d;
```
As shown in the following figure, the location of table **d** is in the specified OBS path.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 10-7 Viewing the location of the d table



- Go to spark-beeline.

  ```
  spark-beeline
  ```

  Create a table.

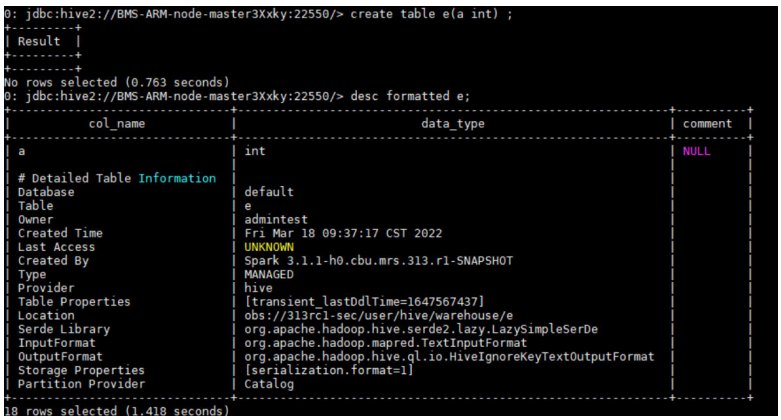
  ```
  create table e(a int);
  ```

  Check the table location.

  ```
  desc formatted e;
  ```

  As shown in the following figure, the location of table **e** is in the specified OBS path.

Figure 10-8 Viewing the location of the e table



Configuring Ranger Permissions

Step 1 Log in to MRS Manager and choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Create a user group without a role, for example, **obs_spark**, and bind the user group to the corresponding user.

Step 3 Log in to the Ranger management page as user **rangeradmin**.

Step 4 On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Click **Add New Policy** to add the **Read** and **Write** permissions on OBS paths to the user group created in **Step 2**. If there are no OBS paths, create one in advance (wildcard character * is not allowed).

----End

NOTE

- To authorize a view chart, you need to grant the view chart permission and the physical table path permission corresponding to the view chart.
- Cascading authorization can be performed only on databases and tables, and cannot be on partitions. If a partition path is not in the table path, you need to manually authorize the partition path.
- Cascading authorization for **Deny Conditions** in the Hive Ranger policy is not supported. That is, the Deny Conditions permission only restricts the table permission and cannot generate the permission of the HDFS/OBS storage source.
- The permission of the HDFS Ranger policy is prior to that of the HDFS/OBS storage source generated by cascading authorization. If the HDFS Ranger permission has been set for the HDFS storage source of the table, the cascading permission does not take effect.
- **alter** operations cannot be performed on tables whose storage source is OBS after cascading authorization. To perform the **alter** operation, you need to grant the **Read** and **Write** permissions of the parent directory of the OBS table path to the corresponding user group.
- Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

Configuring Permissions for CDL Service Users

If Kerberos authentication is enabled for a cluster (security mode) and you want to use CDL to import data to OBS in real time after the cluster is connected to OBS, perform the following steps to grant corresponding users **Read** and **Write** permissions on the corresponding OBS path:

Step 1 Log in to MRS Manager and choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Create a user group with no assigned role, for example, **obs_cdl**, and bind it to the relevant CDL service user, like **cdluser**.

Step 3 Log in to the Ranger management page as user **rangeradmin**.

Step 4 On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

Step 5 Click **Add New Policy** to add the **Read** and **Write** permissions on OBS paths to the created user group. If there are no OBS paths, create one in advance (wildcard character * is not allowed).

The following figure shows the configurations needed for adding the **Read** and **Write** permissions on **obs://***OBS parallel file system name***/cdldata** to user group **obs_cdl**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

10.2.3.8 Accessing OBS Using YARN Through Guardian

After interconnecting Guardian with OBS by referring to **Disabling Ranger OBS Path Authentication for Guardian** or **Enabling Ranger OBS Path Authentication for Guardian**, you can execute YARN jobs on the cluster client to access OBS.

Prerequisites

If you interconnected Guardian with OBS by referring to **Enabling Ranger OBS Path Authentication for Guardian**, ensure that you have the read and write permissions on the OBS path in Ranger. For details about how to obtain the permissions, see **Configuring Ranger Permissions**.

Interconnecting YARN with OBS

Step 1 Log in to the node where the YARN client is installed as the client installation user.

Step 2 Run the following command to switch to the client installation directory.

```
cd Client installation directory
```

Step 3 Run the following command to configure environment variables:

```
source bigdata_env
```

Step 4 If Kerberos authentication is enabled for the cluster, run the following command to perform user authentication. The user must have the read and write permissions on the OBS directory. User authentication is not required for clusters with Kerberos authentication disabled.

```
kinit User performing HDFS operations
```

Step 5 Explicitly add the OBS file system to be accessed in the YARN command line.

- Access the OBS file system.
  ```
  hdfs dfs -ls obs://OBS parallel file system name/path
  ```
- Create a directory in the OBS file system.
  ```
  hdfs dfs -mkdir obs://OBS parallel file system name/hadoop1
  ```
- Execute the YARN task to access OBS.
  ```
  yarn jar Client installation directory/HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-*.jar pi -Dmapreduce.job.hdfs-servers=NAMESERVICE -fs obs://OBS parallel file system name 1 1
  ```

  NAMESERVICE indicates the NameService in HDFS. The default value is **hdfs://hacluster**. If there are multiple NameServices, separate them with commas (**,**).

  Example:

  ```
  yarn jar /opt/hadoopclient/HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-*.jar pi -Dmapreduce.job.hdfs-servers=hdfs://hacluster -fs obs://bucketname 1 1
  ```
- Run the following command to write data to OBS:
  ```
  yarn jar Client installation directory/HDFS/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-*.jar teragen 100  obs://OBS parallel file system name/hadoop1/teragen1
  ```
- Run the following command to copy data from OBS to HDFS:
  ```
  hadoop distcp obs://OBS parallel file system name/hadoop1/teragen1 /tmp
  ```

----End

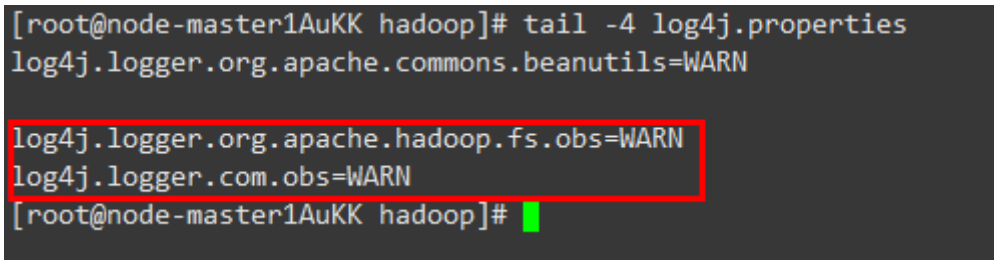
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Changing the Log Level of OBS Client

If a large number of logs are printed in the OBS file system, the read and write performance may be affected. You can adjust the log level of the OBS client as follows:

1. Go to the **hadoop** directory.
   ```
   cd Client installation directory/HDFS/hadoop/etc/hadoop
   ```
2. Edit the file **log4j.properties**.
   ```
   vi  log4j.properties
   ```
   Add the following OBS log level configuration to the file and save it.
   ```
   log4j.logger.org.apache.hadoop.fs.obs=WARN
   log4j.logger.com.obs=WARN
   ```
3. Run the following command:
   ```
   tail -4 log4j.properties
   ```
   If the command output shown in **Figure 10-3** is displayed, the log level has been changed.

   Figure 10-9 Adding an OBS log level

Configuring Ranger Permissions

Step 1 Log in to FusionInsight Manager and choose **System** > **Permission** > **User Group**. On the displayed page, click **Create User Group** to create a user group without any roles, for example, **obs_hadoop1**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

Step 2 Back to FusionInsight Manager and choose **System** > **Permission** > **User**. On the displayed page, click **Create User** to create a user that is associated with the **obs_hadoop1** user group and the **default** role, for example, **hadoopuser1**.

Step 3 Log in to the Ranger management page as the **rangeradmin** user.

Step 4 On the home page, click component plug-in name **OBS** in the **EXTERNAL AUTHORIZATION** area.

Step 5 Click **Add New Policy** and add the **Read** and **Write** permissions on the desired OBS paths to the user group created in **Step 1**.

The following figure shows the configurations needed for adding the **Read** and **Write** permissions on **obs://***OBS parallel file system name***/hadoop1** to user group **obs_hadoop1**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Before configuring permission policies for OBS paths on Ranger, ensure that the AccessLabel function has been enabled for OBS. For how to enable it, contact OBS O&M personnel.

----End

10.3 Interconnecting an MRS Cluster with OBS Using an IAM Agency

10.3.1 Interconnecting an MRS Cluster with OBS Using an IAM Agency

MRS allows you to store data in OBS and use an MRS cluster for data computing only. In this way, storage and compute are decoupled. You can create an IAM agency, which enables ECS to automatically obtain the temporary AK/SK to access OBS. This prevents the permanent AK/SK from being exposed in the configuration file.

By binding an agency, ECSs or BMSs can manage some of your resources. Determine whether to configure an agency based on the actual service scenario. This feature can be used with Hadoop, Hive, Spark, Presto, Sqoop, Flink, and Flume components in clusters. To interconnect MRS with OBS using an IAM agency, perform the following tasks:

1. **Creating an ECS Agency with OBS Access Permissions**
2. **Creating a Decoupled Storage and Compute Cluster**
3. **Creating an OBS Parallel File System for Storing Service Data**
4. **Creating a Lifecycle Rule**

Creating an ECS Agency with OBS Access Permissions

NOTE

- MRS presets **MRS_ECS_DEFAULT_AGENCY** in the IAM agency list by default, allowing you to choose this agency when creating a cluster. This agency has **OBS OperateAccess** permission and, for users with fine-grained policies enabled, **CES FullAccess**, **CES Administrator**, and **KMS Administrator** permissions in the region where the cluster is located. Do not modify **MRS_ECS_DEFAULT_AGENCY** on IAM.
- If you want to use the preset agency, skip the step for creating an agency. If you want to use a custom agency, perform the following steps to create an agency. (To create or modify an agency, you must have the Security Administrator permission.) If you need to have more fine-grained control over the permissions of a specific path in the OBS file system, you can refer to **Configuring Fine- Grained OBS Access Permissions for MRS Cluster Users** to create a custom role policy.

1. Log in to the management console.
2. In the service list, choose **Management & Governance** > **Identity and Access Management**.
3. Choose **Agencies**. On the displayed page, click **Create Agency**.
4. Set **Agency Name**. For example, enter **mrs_ecs_obs**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

5. Set **Agency Type** to **Cloud service** and select **ECS BMS** to authorize ECS or BMS to invoke OBS.
6. Set **Validity Period** to **Unlimited** and click **Done**.
7. In the displayed dialog box, click **Authorize**. Search for **OBS OperateAccess** and select it.
8. Click **Next**. On the page that is displayed, select the desired scope for the permissions you selected. By default, **All resources** is selected. Click **Show More**, select **Global resources**, and click **OK**.
9. In the dialog box that is displayed, click **OK** to start authorization. After the message "**Authorization successful.**" is displayed, click **Finish**. The agency is successfully created.

Creating a Decoupled Storage and Compute Cluster

You can configure an agency when creating a cluster or bind an agency to an existing cluster to decouple storage and compute. This section uses a cluster with Kerberos authentication enabled as an example.

Configuring an agency when creating a cluster:

1. Log in to the MRS console.
2. Click **Create Cluster**. The page for creating a cluster is displayed.
3. Click the **Custom Config** tab.
4. On the Custom Config page, set cluster parameters based on your service requirements.

Table 10-3 Parameters for a custom MRS cluster

Parameter	Description	Example Value
Region	Region where the resource to be created is located. Resources in different regions cannot communicate with each other over the intranet. To improve access speed and reduce network latency, select the region nearest to you.	-
Cluster Name	MRS cluster name. You can use the default name. However, you are advised to include a project name abbreviation or date in the cluster name for remembering and distinguishing. After a cluster is created, you can change the cluster name in the cluster list.	mrs-test

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Cluster Type	Select a required MRS cluster type based on service requirements. • Analysis cluster: It is well-suited for analyzing and processing large amounts of offline data to obtain results. Recommended components include Hadoop, Spark, Hive, Flink, Oozie, and Tez. • Streaming cluster: It is ideal for streaming data processing and quick analysis of real-time data sources. Recommended components include Kafka and Flume. • Hybrid cluster: It is suitable for both offline data analysis and stream processing. • Custom: A wide range of components are available. You can select supported components as required based on your MRS cluster version.	Custom
Version Type	MRS provides two cluster version types: **LTS** and **Normal**. Different version types provide different components. You can select a version type as required. • **LTS**: employs MRS's own components to provide highly reliable clusters with strong DR capabilities, making long-term support and evolution possible. • **Normal**: integrates MRS's mature and stable features and functions with open-source capabilities, offering high performance and stability.	LTS
Cluster Version	Version of the MRS cluster. Different versions may contain different open-source component versions and functions. You are advised to select the latest version.	MRS 3.2.0-LTS.1

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Component	Cluster templates containing preset open-source components you will need for your business. Components that can connect to OBS through IAM agencies include Hadoop, Hive, Spark, Presto, Sqoop, Flink, and Flume. Select the components based on your requirements. **NOTE** Hadoop components include HDFS, YARN, and MapReduce. Components used within a cluster, such as DBService, KrbServer, and LdapServer, are not displayed in the component list during cluster creation.	Hadoop and Hive
Metadata	Whether to use external data sources to store Hive and Ranger metadata of the cluster. • **Local**: Metadata is stored in the local cluster. • **External data connection**: Metadata of external data sources is used. If the cluster is abnormal or deleted, metadata is not affected. This mode applies to scenarios where storage and compute are decoupled.	Local
AZ	AZ to which resources to be created belong in the current region. An AZ is a physical area where resources use independent power supply and networks.	AZ1
VPC	VPC to which the MRS cluster node belongs. If no VPC is available, click **View VPC** to access the network console and create a VPC.	-
Subnet	Subnet information in the VPC. If no subnet is available, click **View Subnet** to access the network console and create a subnet.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Security Group	A security group is a set of ECS access rules. It provides access policies for ECSs that have the same security protection requirements and are mutually trusted in a VPC. When you create an MRS cluster, a security group is automatically created by default. You can also select an existing security group from the drop-down list.	Auto create
Common Node Configurations	This parameter is available only when **Cluster Type** is set to **Custom**. **Common Node Configurations** can be **Compact**, **Full-size**, or **OMS-separate**.	Compact
Cluster Nodes	Specifications and quantity of nodes in an MRS cluster. For MRS 3.*x* or later, the memory of a Master node must be greater than 64 GB.	Set the number of cluster nodes and node specifications as required.
Kerberos Authentication	Whether to enable Kerberos authentication for each component in the MRS cluster. If Kerberos authentication is enabled, users can access component resources only after being authenticated. This option cannot be changed after you create a cluster.	Kerberos authentication enabled
Username	Default user for logging in to MRS Manager and nodes in the MRS cluster. User **admin** is used to log in to MRS Manager, while user **root** is the OS user used to log in to the nodes in the cluster.	-

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Password/ Confirm Password	Password of the MRS Manager administrator **admin**. Keep the password secure. ● Must contain 8 to 26 characters. ● Must contain every type of the following: – Lowercase letters – Uppercase letters – Digits – Special characters: `~!@#$ %^&*()-_=+|[{}];:',<.>/? ● Cannot be the same as the username or the username spelled backwards.	-
Login Mode	Mode for logging in to a node in the MRS cluster. ● Password Use a password to log in to the node as user **root**. You need to set the password of user **root**. ● Key Pair Use a key pair to log in to the node. Select a key pair from the drop-down list. If you have obtained the private key file, select **I acknowledge that I have obtained private key file** *SSHkey-xxx* **and that without this file I will not be able to log in to my ECS.** If no key pair is available, click **View Key Pair** to create or import a key pair, and then obtain the private key file.	Password

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Advanced Configuration	Enable **Advanced Configuration** and set **Agency**. After an agency is bound, the ECS or BMS services will have the permission to manage some resources in the cluster. • Click **Available agencies** and select the agency created in **Creating an ECS Agency with OBS Access Permissions** from the drop-down list. • Select the **MRS_ECS_DEFAULT_AGENCY** agency preset by MRS in IAM.	MRS_ECS_DEFAULT_AGENCY
Enterprise Project	An enterprise project facilitates project-level management and grouping of cloud resources and users. You can select the system-defined enterprise project **default** or create a new one.	default
Secure Communications	To allow the MRS console to access big data components in your VPC, you need to enable relevant security group rules to allow traffic to pass. For details, see **Configuring Secure Communication Authorization for an MRS Cluster**.	Select the check box.

5. Click **Apply Now** and wait until the cluster is created.

Configuring an agency for an existing cluster:

1. Log in to the MRS console. In the navigation pane on the left, choose **Active Clusters**.
2. Click the name of the cluster to enter its details page.
3. On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync** to synchronize IAM users.
4. On the **Dashboard** tab page, click **Select Agency** on the right side of **Agency** to select an agency and click **OK** to bind it. Alternatively, click **Create New** to go to the IAM console to create an agency and select it.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Creating an OBS Parallel File System for Storing Service Data

> **WARNING**
>
> In decoupled storage-compute scenarios, make sure to use an OBS parallel file system. Using an OBS bucket can significantly affect the performance of the cluster.

1. Log in to the OBS console.
2. Choose **Parallel File Systems** > **Create Parallel File System**.
3. Enter the file system name, for example, **mrs-word001**.

   Set other parameters as required.
4. Click **Create Now**.
5. In the parallel file system list on the OBS console, click the file system name to go to the details page.
6. In the navigation pane, choose **Files** and create the **program** and **input** folders.
   - **program**: Upload the program package to this folder.
   - **input**: Upload the input data to this folder.

Creating a Lifecycle Rule

In MRS 3.2.0-LTS.1 and later versions, components prevent mis-deletion by default. That is, file data deleted by component users is not directly deleted but stored in the recycle bin directory in the OBS file system.

To save OBS space, you need to enable periodical deletion of file data from the OBS recycle bin by referring to **Configuring the Policy for Clearing Recycle Bin Directories of MRS Cluster Components**.

Managing MRS Cluster Agencies

In practical scenarios, if the agency currently bound to the cluster cannot meet the requirements, you can switch the agency by performing the following operations:

1. Log in to the MRS console and click the name of the desired MRS cluster in the active cluster list.
2. In the **O&M Management** area, click **Select Agency** on the right of **Agency**.
3. In the displayed **Select Agency** dialog box, select the agency you want to switch to and click **OK**. Alternatively, click **Create New** to open the IAM console, create a new agency, and bind it here.
4. After switching the agency, log in to MRS Manager by referring to **Accessing MRS Manager**.
5. Choose **Cluster** > **Services** > **Meta**. In the upper right corner of the page, click **More** and select **Restart Service**. Enter the password of the current user and click **OK** to restart the meta service for the new agency to take effect.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTICE

Due to the IAM agency permission cache mechanism, changes to agency permissions take effect within 24 hours. If you need the permission change to take effect immediately, it is recommended that you create a new agency with the required permissions and bind it to the MRS cluster.

10.3.2 Configuring the Policy for Clearing Recycle Bin Directories of MRS Cluster Components

Scenarios

In MRS 3.2.0-LTS.1 or later, components prevent mis-deletion by default. That is, file data deleted by component users is not directly deleted but stored in the recycle bin directory in the OBS file system. This function is compatible with the native garbage collection mechanism of Hadoop FS, providing additional data protection for OBS-based Hadoop big data systems.

This section describes how to set the lifecycle policy of the recycle bin directory in the OBS file system to automatically clear related data periodically. The recycle bin directory is created for each required user. If a new user is added to the MRS cluster and has the permission to delete component data, you need to configure the policy for clearing the recycle bin directory for the new user.

CAUTION

For clusters that use decoupled storage and compute, configure a lifecycle policy for the related directories by referring to this section. Otherwise, the storage space may be used up and storage fees may increase.

You need to configure lifecycle policies for the recycle bin directories of preset users in the MRS cluster and the recycle bin directories of new users who need accidental deletion prevention. If a low privileged agency is used or only the permission for MRS users to access OBS file system directories is configured by referring to **Configuring Fine- Grained OBS Access Permissions for MRS Cluster Users**, you will need the operation permission for the recycle bin directory.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 10-4 Directories for which a lifecycle policy needs to be configured

Cluster Version	Directory Type	Component	Directory	How to Create
Versions earlier than MRS 3.3.0-LTS	Recycle bin directories that must be configured by default for each component in an MRS cluster	Hive	• user/omm/.Trash • user/hive/.Trash	If the **.Trash** folder does not exist, create it on the cluster client as user **omm**. Run the following command: **hdfs dfs -mkdir -p obs://***Name of the OBS parallel file system where the table is stored***/***Folder path*
		Spark	• user/omm/.Trash • user/root/.Trash • user/spark2x/.Trash	
		HetuEngine	• user/omm/.Trash • user/hetuserver/.Trash	
		HBase	• user/hbase/.Trash • user/omm/.Trash	
	Recycle bin directories of users who need accidental deletion prevention	Hive/Spark/HetuEngine	user/*<New service user>*/.Trash	
MRS 3.3.0-LTS or later	Default recycle bin directories configured for each component in an MRS cluster	Hive/Spark/HetuEngine	/user/.Trash	

For example, if a user with the following permissions has been added to the cluster, you need to create a recycle bin directory clearing rule for the user in the parallel file system:

- Permissions to delete the HDFS files
- **DROP**, **INSERT OVERWRITE**, and **TRUNCATE** permissions on Hive tables
- **DROP**, **TRUNCATE**, **DELETE**, **INSERT OVERWRITE**, and **LOAD OVERWRITE** permissions on HetuEngine

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

The following table lists the MRS component commands that support data restoration after accidental deletion.

Table 10-5 Commands supporting data restoration after accidental deletion

Component	Commands Supporting Data Restoration After Accidental Deletion
Hive	• Deleting a table: drop table • Clearing table data: truncate table • Deleting a table partition: alter table drop partition • Overwriting a table or partition: insert overwrite table [partition] • Overwriting a directory: insert overwrite directory • Overwriting a table or partition: load data overwrite
Spark	• Deleting a table: drop table • Clearing table data: truncate table • Clearing a table partition: truncate table table_name partition • Deleting a table partition: alter table drop partition • Overwriting a table or partition: insert overwrite table [partition] • Overwriting a directory: insert overwrite directory • Overwriting a table or partition: load data overwrite
HetuEngine	• Deleting a table: drop table • Clearing table data: truncate table delete table • Deleting a table partition: alter table drop partition delete table where [partition] • Overwriting a table or partition: insert overwrite table [partition] • Overwriting a table or partition: load data overwrite

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Component	Commands Supporting Data Restoration After Accidental Deletion

Hudi	• Deleting a table:
	`drop table`
	• Deleting a table partition:
	– SQL:
	`alter table drop partition`
	– API:
	`delete_partition`
	• Deleting row-level data:
	– SQL:
	`delete from`
	– API:
	`delete`
	• Clearing table data:
	`truncate table`
	• Overwriting a table:
	– SQL:
	`insert overwrite`
	`insert overwrite table`
	– API:
	`insert_overwrite`
	`insert_overwrite_table`

Configuring the Lifecycle Rule of an OBS Directory

Step 1 Log in to the OBS console.

Step 2 Click **Parallel File Systems** and click the name of the file system used by the current MRS cluster.

Step 3 In the navigation pane, choose **Data Management** > **Lifecycle Rules**. Click **Create** to create a lifecycle rule for a specified directory.

Table 10-6 Parameters for creating a lifecycle rule

Parameter	Description	Example Value
Status	Whether to enable the lifecycle rule.	Enable
Rule Name	Rule name that identifies different lifecycle configurations.	rule-test

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description	Example Value
Prefix	Prefix of the objects to which the lifecycle rule applies. Objects that have the specified prefix will be managed by the lifecycle rule. The prefix cannot start with a slash (/), have consecutive slashes (/), or contain the following special characters: **\:*?"<>|** If this parameter is not specified, the rule will take effect for the entire file system. **WARNING** To prevent other service data from being deleted by mistake, you are not advised to use the lifecycle rule configured for the entire file system or high-level directories. Generally, the recycle bin directory of MRS components is in the following format. If the folder does not exist, create it. user/*<Username>*/.Trash	user/omm/.Trash
Delete Files After (Days)	The object within the rule configuration scope expires and is automatically deleted by OBS if the number of days since its last update reaches this parameter value. The recommended value is 1 to 7 days.	2 days

Step 4 Click **OK** to complete the lifecycle rule configuration.

You can click **Edit** in the **Operation** column of a lifecycle rule to edit it. You can also click **Disable** or **Enable** to disable or enable it.

Step 5 Repeat the preceding steps to create recycle bin directory clearing rules for all users who have the data deletion permission in the current MRS cluster one by one until all recycle bin directories in the OBS file system are configured.

----End

10.3.3 Example for Interconnecting an MRS Cluster with OBS

10.3.3.1 Interconnecting Flink with OBS Using an IAM Agency

After configuring a cluster with decoupled storage and compute by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can access the OBS parallel file system and run jobs in a Flink task. This section describes how to submit a Flink WordCount job.

Prerequisites

Before interconnecting Flink with OBS, ensure that YARN is connected to OBS as Flink jobs run on YARN.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Interconnecting Flink with OBS

Step 1 Log in to the Flink client installation node as the client installation user.

Step 2 Initialize environment variables.

```
source Client installation directory/bigdata_env
```

Step 3 Configure the Flink client.

Step 4 Start a session.

- Normal cluster (Kerberos authentication disabled)

  ```
  yarn-session.sh -nm "session-name" -d
  ```

- Security cluster (Kerberos authentication enabled)
  - If the paths of the **flink.keystore** and **flink.truststore** files are relative ones:

    Run the following command in the directory at the same level as **ssl** to start the session. **ssl/** is a relative path.

    Go to the **conf** directory.

    ```
    cd Client installation directory/Flink/flink/conf
    ```

    Start a session.

    ```
    yarn-session.sh -t ssl/ -nm "session-name" -d
    ```

    The session is started successfully.

    ```
    ...
    Cluster started: Yarn cluster with application id application_1624937999496_0017
    JobManager Web Interface: http://192.168.1.150:32261
    ```

  - If the **flink.keystore** and **flink.truststore** files are stored in absolute paths, run the following commands to start a session:

    Go to the **conf** directory.

    ```
    cd Client installation directory/Flink/flink/conf
    ```

    Start a session.

    ```
    yarn-session.sh -nm "session-name" -d
    ```

Step 5 Authenticate the user of the cluster with Kerberos authentication enabled. Skip this step for the user of the cluster with Kerberos authentication disabled.

```
kinit Username
```

Step 6 Add the target OBS file system to the Flink CLI and execute the analysis program.

1. Create the **test** file in the **/tmp/test** directory.

   ```
   echo -e 'test' >/tmp/test
   ```

2. Create a directory for storing the **test** file in the OBS parallel file system.

   ```
   hdfs dfs -mkdir -p obs://OBS parallel file system name/tmp/flinkjob
   ```

3. Upload the **test** file to the OBS parallel file system.

   ```
   hdfs dfs -put /tmp/test/ obs://OBS parallel file system name/tmp/flinkjob/
   ```

4. Run **WordCount.jar**.

   ```
   flink run Client installation directory/Flink/flink/examples/batch/WordCount.jar -input obs://OBS parallel file system name/tmp/flinkjob/test -output obs://OBS parallel file system name/tmp/flinkjob/output
   ```

Step 7 Log in to the OBS console and view the execution result of **WordCount.jar** in the output path specified by **Step 6.4**.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

10.3.3.2 Interconnecting Flume with OBS Using an IAM Agency

After configuring decoupled storage and compute for a cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can run OBS jobs using Flume. This section describes how to submit a Flume job that reads a **.txt** file from the client node and exports the file to the OBS directory.

Notes and Constraints

This section applies to MRS 3.*x* or later.

Interconnecting Flume with OBS

Step 1 Create an OBS folder for storing data.

1. Log in to the OBS console.
2. In the navigation pane on the left, choose **Resources** > **Parallel File Systems**.
3. On the displayed page, click the name of the parallel file system you created to access its details page.
4. In the navigation pane on the left, choose **Files**. On the displayed page, click **Create Folder** to create the **testFlumeOutput** folder.

Step 2 Log in to the node where the Flume client is installed as user **root**.

Step 3 Create the **/opt/flumeInput** directory and create a customized **.txt** file in it.

Step 4 Add the following content to the *Client installation directory*/**FusionInsight-flume-*/properties.properties** file:

```
# source
server.sources = r1
# channels
server.channels = c1
# sink
server.sinks = obs_sink
# ----- define net source -----
server.sources.r1.type = seq
server.sources.r1.spooldir = /opt/flumeInput
# ---- define OBS sink ----
server.sinks.obs_sink.type = hdfs
server.sinks.obs_sink.hdfs.path = obs://esdk-c-test-pfs1/testFlumeOutput
server.sinks.obs_sink.hdfs.filePrefix = %[localhost]
server.sinks.obs_sink.hdfs.useLocalTimeStamp = true
# set file size to trigger roll
server.sinks.obs_sink.hdfs.rollSize = 0
server.sinks.obs_sink.hdfs.rollCount = 0
server.sinks.obs_sink.hdfs.rollInterval = 5
#server.sinks.obs_sink.hdfs.threadsPoolSize = 30
server.sinks.obs_sink.hdfs.fileType = DataStream
server.sinks.obs_sink.hdfs.writeFormat = Text
server.sinks.obs_sink.hdfs.fileCloseByEndEvent = false

# define channel
server.channels.c1.type = memory
server.channels.c1.capacity = 1000
# transaction size
server.channels.c1.transactionCapacity = 1000
server.channels.c1.byteCapacity = 800000
server.channels.c1.byteCapacityBufferPercentage = 20
server.channels.c1.keep-alive = 60
server.sources.r1.channels = c1
server.sinks.obs_sink.channel = c1
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Set **server.sources.r1.spooldir** to the directory of the **.txt** file created in **Step 3**.
- Set **server.sinks.obs_sink.hdfs.path** to the OBS file system created in **Step 1**.

Step 5 Copy the **hadoop-huaweicloud-*.jar** and **mrs-obs-provider-*.jar** files from *Client installation directory*/**Hive/Beeline/lib** to *Flume client installation directory*/**fusionInsight-flume-*/lib**. Then run the following commands to modify permissions:

1. Go to the **lib** directory.
   cd *Flume client installation directory*/fusionInsight-flume-*/lib
2. Modify the permission on **hadoop-huaweicloud-*.jar**.
   chmod 755 hadoop-huaweicloud-*.jar
3. Modify the permission on **mrs-obs-provider-*.jar**.
   chmod 755 mrs-obs-provider-*.jar

Step 6 Run the following command to restart the Flume client:

Go to the **bin** directory.

cd *Flume client installation directory*/fusionInsight-flume-*/bin

Restart the Flume client.

./flume-manager.sh restart

Step 7 View the result in the OBS system.

1. Log in to the OBS console.
2. In the navigation pane on the left, choose **Resources** > **Parallel File Systems**. Click the name of the parallel file system you created. In the navigation pane on the left, choose **Files**. On the displayed page, click the folder created in **Step 1** to view the result.

----End

10.3.3.3 Interconnecting HDFS with OBS Using an IAM Agency

After configuring decoupled storage and compute for a cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can view and create OBS file directories on the HDFS client.

Interconnecting HDFS with OBS

Step 1 Log in to the node where the HDFS client is installed as the client installation user.

For details about how to download and install the cluster client, see .

Step 2 Run the following command to switch to the client installation directory.
cd *Client installation directory*

Step 3 Run the following command to configure environment variables:
source bigdata_env

Step 4 Authenticate the user of the cluster with Kerberos authentication enabled. Skip this step for clusters with Kerberos authentication disabled.
kinit *Component service user*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Add the target OBS file system in the HDFS CLI.

For example:

- Run the following command to access the OBS file system:
  ```
  hdfs dfs -ls obs://OBS parallel file system name/path
  ```
  For example, run the following command to access the **mrs-word001** parallel file system. If the file list is returned, OBS is successfully accessed.
  ```
  hadoop fs -ls obs://mrs-word001/
  ```

  Figure 10-10 Returned file list

  ```
  Found 2 items
  drwxrwxrwx   - root root          0 2019-12-21 11:04 obs://mrs-word001/input
  drwxrwxrwx   - root root          0 2019-12-21 11:04 obs://mrs-word001/program
  ```

- Run the following command to upload the **/opt/test.txt** file from the client node to the OBS file system path:
  ```
  hdfs dfs -put /opt/test.txt obs://OBS parallel file system name/path
  ```
  View the uploaded file in the corresponding OBS file system path.

----End

Changing the Log Level of OBS Client

If a large number of logs are printed in the OBS file system, the read and write performance may be affected. You can adjust the log level of the OBS client as follows:

1. Go to the **hadoop** directory.
   ```
   cd Client installation directory/HDFS/hadoop/etc/hadoop
   ```
2. Edit the file **log4j.properties**.
   ```
   vi  log4j.properties
   ```
   Add the following OBS log level configuration to the file and save it.
   ```
   log4j.logger.org.apache.hadoop.fs.obs=WARN
   log4j.logger.com.obs=WARN
   ```
3. Run the following command:
   ```
   tail -4 log4j.properties
   ```
   If the command output shown in **Figure 10-11** is displayed, the log level has been changed.

   Figure 10-11 Viewing the log level

   ```
   [root@ecs-          hadoop]# tail -4 log4j.properties
   # Log levels of third-party libraries
   log4j.logger.org.apache.commons.beanutils=WARN
   log4j.logger.org.apache.hadoop.fs.obs=WARN
   log4j.logger.com.obs=WARN
   [root@ecs-          hadoop]#
   ```

10.3.3.4 Interconnecting Hive with OBS Using an IAM Agency

After configuring decoupled storage and compute for a cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can create tables with OBS paths as their location on the Hive client.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Setting the Location to an OBS Path When Creating a Table

Step 1 Log in to the client installation node as the client installation user.

Step 2 Run the following command to initialize environment variables:

```
source Client installation directory/bigdata_env
```

Step 3 Authenticate the user of the cluster with Kerberos authentication enabled. (The user must have the permission to perform operations on Hive.) Skip this step for the user of the cluster with Kerberos authentication disabled.

```
kinit User performing Hive operations
```

Step 4 Log in to FusionInsight Manager of a cluster earlier than MRS 3.2.0-LTS, choose **Cluster** > **Services** > **Hive**, and click **Configurations** then **All Configurations**.

In the navigation pane on the left, choose **Hive(Service)** > **Customization**. In the custom parameters, add **dfs.namenode.acls.enabled** to **hdfs.site.customized.configs** and set its value to **false** to disable the HDFS ACL function.

Step 5 Click **Save** to save the configuration for versions earlier than MRS 3.2.0-LTS. On the **Dashboard** page, click **More** and select **Restart Service**. Enter the password of the current user, click **OK**, and select **Restart upper-layer services**. Click **OK** to restart Hive.

Step 6 Log in to the beeline client and set **Location** to the OBS file system path when creating a table.

```
beeline
```

For example, run the following command to create the table **test** in **obs://***OBS parallel file system name***/user/hive/warehouse/***Database name***/***Table name*:

```
create table test(name string) location "obs://OBS parallel file system name/user/hive/warehouse/Database name/Table name";
```

Check the table location.

```
desc formatted test;
```

NOTE

You need to attach the component operation user to the URL policy within the Ranger policy. Enter the complete OBS path of the object for **URL** and select **Read** and **Write** permissions. Other permissions are not involved in the URL policy.

----End

Interconnecting Hive with OBS Through MetaStore

Step 1 Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Hive** > **Configurations** > **All Configurations**.

For details about how to log in to MRS Manager, see **Accessing MRS Manager**.

- For versions earlier than MRS 3.2.0-LTS:
  - In the navigation pane on the left, choose **MetaStore(role)** > **Customization**. Add the configuration item **hive.metastore.warehouse.dir** to the custom parameter **hive.metastore.customized.configs** and set the value to an OBS path,

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

for example, **obs://hivetest/user/hive/warehouse/**, where **hivetest** is the name of the OBS parallel file system.

- In the navigation pane on the left, choose **HiveServer(role)** > **Customization**. Add the configuration item **hive.metastore.warehouse.dir** to custom parameters **hive.metastore.customized.configs** and **hive.server.customized.configs** and set the values to an OBS path, for example, **obs://hivetest/user/hive/warehouse/**, where **hivetest** is the name of the OBS parallel file system.

- For MRS 3.2.0-LTS and later versions:

  Search for **hive.metastore.warehouse.dir** in the search box and change its value to an OBS path, for example, **obs://hivetest/user/hive/warehouse/**, where **hivetest** indicates the OBS file system name.

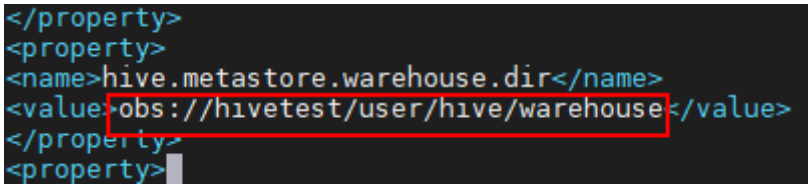
Step 2 Save the change and restart Hive.

Step 3 (Optional) Install the client by referring to **Installing an MRS Cluster Client**. If the client has been installed in the cluster, go to **Step 4**.

Step 4 Update the client configuration file.

1. Run the following command to open **hivemetastore-site.xml** in the Hive configuration file directory on the client:
   ```
   vim Client installation directory/Hive/config/hivemetastore-site.xml
   ```
2. Change the value of **hive.metastore.warehouse.dir** to the corresponding OBS path, for example, **obs://hivetest/user/hive/warehouse/**, where **hivetest** is the OBS bucket name.

   Figure 10-12 Configuring the OBS Path

3. For MRS 3.2.0-LTS and later versions, change the value of **hive.metastore.warehouse.dir** in **hivemetastore-site.xml** file to the corresponding OBS path, for example, **obs://hivetest/user/hive/warehouse/**. The XML file is stored in the HCatalog client configuration file directory.
   ```
   vi Client installation directory/Hive/HCatalog/conf/hivemetastore-site.xml
   ```

Step 5 Log in to the beeline client, create a table, and check whether the location is the OBS path.

1. Go to the client installation directory.
   ```
   cd Client installation directory
   ```
2. Load the environment variables.
   ```
   source bigdata_env
   ```
3. Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.
   ```
   kinit User performing Hive operations
   ```
4. Log in to the Hive client.
   ```
   beeline
   ```
5. Create the **test** table.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

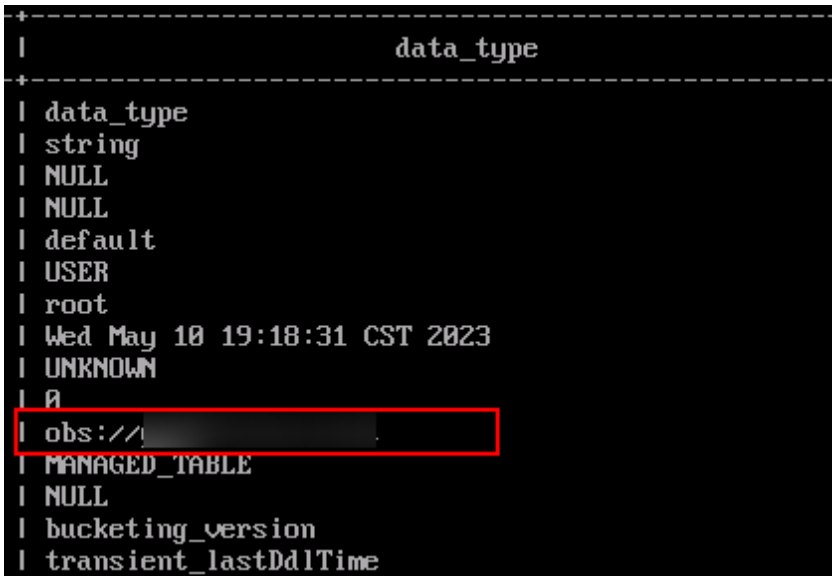
```
create table test(name string);
```

6. Check the table location.

```
desc formatted test;
```

Location of the table is the OBS path.

Figure 10-13 Viewing the location of a Hive table

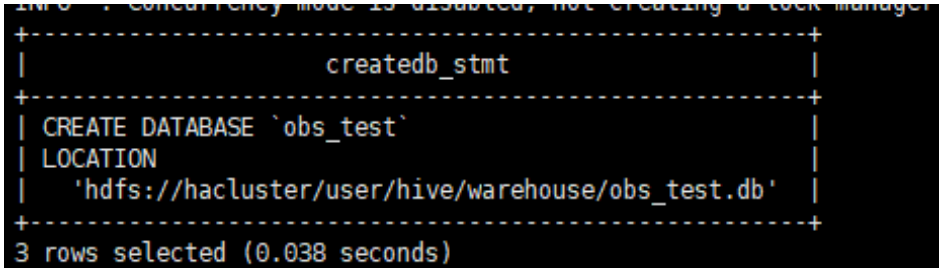


Step 6 If the location of the current database points to HDFS, tables created in the database are located in HDFS by default. You do not need to specify the location. To modify the default table creation policy, update the location of the database to OBS. The procedure is as follows:

1. Query the location of the database.

```
show create database obs_test;
```

Figure 10-14 Viewing the location of a Hive table



2. Change the database location.

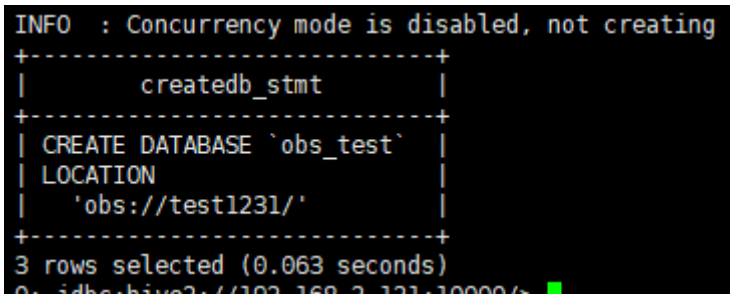
```
alter database obs_test set location 'obs://OBS parallel file system name/user/hive/warehouse/Database name';
```

Check whether the database location is changed to an OBS path.

```
show create database obs_test;
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 10-15 Checking the new location of the Hive table



3. Change the table location.

   alter table *user_info* set location 'obs://*OBS parallel file system name/user/hive/warehouse/Database name/Table name*';

   If there is already data in the table, migrate the data to the new location.

----End

10.3.3.5 Interconnecting Hudi with OBS Using an IAM Agency

After configuring decoupled storage and compute for a cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can create Hudi COW tables in spark-shell and store them to OBS.

Interconnecting Hudi with OBS

Step 1 Log in to the client installation node as the client installation user.

Step 2 Run the following commands to configure environment variables:

Load the environment variables.

```
source Client installation directory/bigdata_env
```

Load the component environment variables.

```
source Client installation directory/Hudi/component_env
```

Step 3 Modify the configuration file:

```
vim Client installation directory/Hudi/hudi/conf/hdfs-site.xml
```

Modify the following content, where the **dfs.namenode.acls.enabled** parameter specifies whether to enable the HDFS ACL function.

```
<property>
<name>dfs.namenode.acls.enabled</name>
<value>false</value>
</property>
```

Step 4 Authenticate the user of the cluster with Kerberos authentication enabled. Skip this step for the user of the cluster with Kerberos authentication disabled.

```
kinit Username
```

Step 5 Start spark-shell and run the following commands to create a COW table and save it in OBS:

```
import org.apache.hudi.QuickstartUtils._
import scala.collection.JavaConversions._
import org.apache.spark.sql.SaveMode._
import org.apache.hudi.DataSourceReadOptions._
import org.apache.hudi.DataSourceWriteOptions._
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
import org.apache.hudi.config.HoodieWriteConfig._
val tableName = "hudi_cow_table"
val basePath = "obs://testhudi/cow_table/"
val dataGen = new DataGenerator
val inserts = convertToStringList(dataGen.generateInserts(10))
val df = spark.read.json(spark.sparkContext.parallelize(inserts, 2))
df.write.format("org.apache.hudi").
options(getQuickstartWriteConfigs).
option(PRECOMBINE_FIELD_OPT_KEY, "ts").
option(RECORDKEY_FIELD_OPT_KEY, "uuid").
option(PARTITIONPATH_FIELD_OPT_KEY, "partitionpath").
option(TABLE_NAME, tableName).
mode(Overwrite).
save(basePath);
```

In the preceding command, **obs://testhudi/cow_table/** indicates the OBS path, and **testhudi** indicates the name of the parallel file system. Change them as required.

Step 6 Use DataSource to check whether the table is created and whether the data is normal.

```
val roViewDF = spark.
read.
format("org.apache.hudi").
load(basePath + "/*/*/*/*")
roViewDF.createOrReplaceTempView("hudi_ro_table")
spark.sql("select * from  hudi_ro_table").show()
```

Figure 10-16 Viewing table data

```
+-------------------+--------------------+--------------------+----------------------+--------------------+--------------------+---------+-------
-------------------+----------+--------------------+--------------------+--------------------+--------------------+---------+--------
-----+--------------------+
|_hoodie_commit_time|_hoodie_commit_seqno|  _hoodie_record_key|_hoodie_partition_path|   _hoodie_file_name|           begin_lat|
          begin_lon|              driver|             end_lat|             end_lon|                fare|       partitionpath|    rider|
  ts|                uuid|
+-------------------+--------------------+--------------------+----------------------+--------------------+--------------------+---------+-------
-------------------+----------+--------------------+--------------------+--------------------+--------------------+---------+--------
-----+--------------------+
|     20230516204009|  20230516204009_1_3|4631e932-0c11-40e...|  americas/united_s...|ldd9a945e-fdd6-418...|0.21624150367601136|0
.14285851259466197|driver-213| 0.5890949624813784| 0.09668230319271115| 93.56018115236618|americas/united_s...|rider-213|16842065
90937|4631e932-0c11-40e...|
|     20230516204009|  20230516204009_1_6|4946e241-5f5d-4b0...|  americas/united_s...|ldd9a945e-fdd6-418...|0.11488393157088261|
0.6273212202489661|driver-213| 0.7454678537511295| 0.39549390649008973| 27.794786085825961americas/united_s...|rider-213|16838202
22432|4946e241-5f5d-4b0...|
|     20230516204009|  20230516204009_1_8|0c0f584a-d8d5-48c...|  americas/united_s...|ldd9a945e-fdd6-418...| 0.8742041526408587|
0.7528268153249502|driver-213| 0.9197827128888302|  0.3624647708744041|19.179139106643607|americas/united_s...|rider-213|16841648
3501|0c0f584a-d8d5-48c...|
|     20230516204009|  20230516204009_1_9|a8609706-0af8-476...|  americas/united_s...|ldd9a945e-fdd6-418...| 0.5731835407930634|
0.4923479652912024|driver-213|0.8098858178093026160.42520899698713666| 64.27696295884016|americas/united_s...|rider-213|16836812
23848|a8609706-0af8-476...|
|     20230516204009|  20230516204009_1_10|2fb4e48e-18a5-4d7...|  americas/united_s...|ldd9a945e-fdd6-418...| 0.1856488085068272|
0.9694586417048392|driver-213|0.38186367037201974|0.25252652214479043| 33.922164839486431americas/united_s...|rider-213|16841375
36499|2fb4e48e-18a5-4d7...|
|     20230516204009|  20230516204009_0_2|eb27db8e-d5a3-486...|  americas/brazil/s...|0e7bd523-bbed-4c9...| 0.4726905879569653|0
.46157858450465483|driver-213|  0.7548034070008858| 0.9671159942018241|34.158284716382045|americas/brazil/s...|rider-213|16840336
61640|eb27db8e-d5a3-486...|
|     20230516204009|  20230516204009_0_4|18d28fa8-3290-439...|  americas/brazil/s...|0e7bd523-bbed-4c9...| 0.0750588760043035|0
.03044104444445928|driver-213|0.04376353354538354| 0.6346040067610669| 66.62804366450246|americas/brazil/s...|rider-213|16838970
19204|18d28fa8-3290-439...|
|     20230516204009|  20230516204009_0_7|ac77c162-6d0c-446...|  americas/brazil/s...|0e7bd523-bbed-4c9...| 0.6100070562136587|
0.8779402295427752|driver-213| 0.3407070505929602| 0.5030798142293655|  43.492381121901412americas/brazil/s...|rider-213|16840370
60550|ac77c162-6d0c-446...|
|     20230516204009|  20230516204009_2_1|9270e4c6-9d50-482...|    asia/india/chennai|9e649e53-b5ef-4bf...|  0.406135109773087|
0.5644092139040959|driver-213|  0.7987063049415171|0.026983592271828341|17.851135255091155|  asia/india/chennai|rider-213|16841939
61814|9270e4c6-9d50-482...|
|     20230516204009|  20230516204009_2_5|44a595df-e111-409...|    asia/india/chennai|9e649e53-b5ef-4bf...|  0.6510585056607421
0.8192868687714224|driver-213|0.20714896002914462|0.06224031095026987| 41.062909290463681  asia/india/chennai|rider-213|16842167
33290|44a595df-e111-409...|
+-------------------+--------------------+--------------------+----------------------+--------------------+--------------------+---------+-------
-------------------+----------+--------------------+--------------------+--------------------+--------------------+---------+--------
```

Step 7 Exit the **spark-shell** CLI.

```
:q
```

----End

10.3.3.6 Interconnecting MapReduce with OBS Using an IAM Agency

After configuring a storage-compute decoupled cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can add a Mapreduce custom configuration to enable JobHistory to delete log files of the user directory in a timely manner.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Interconnecting MapReduce with OBS

Step 1 Log in to the MRS management console. In the navigation pane on the left, choose **Clusters** > **Active Clusters**. On the displayed page, click the name of the cluster you created to access its details page.

Step 2 Choose **Components** > **MapReduce**.On the displayed page, click **Configurations** then **All Configurations**. In the navigation pane on the left, choose **MapReduce** > **Customization**. In the custom configuration items, add the configuration item **mapreduce.jobhistory.always-scan-user-dir** to **core-site.xml** file.

This parameter specifies whether the JobHistory service always scans log files in the user directory when scanning job logs. The value can be **true** or **false**.

- **true**: The JobHistory service forcibly scans the specified directory even if the directory has been scanned. This ensures that all related log files are correctly collected and processed each time a job is executed.
- **false** (default value): The JobHistory service does not scan the user directories that have been scanned. This reduces unnecessary disk I/O operations.

Set this parameter to **true**.

Step 3 Click **Save Configuration**.

Step 4 Click the **Service Status** tab and choose **More** > **Restart Service** to restart the MapReduce service.

Step 5 If the OBS job is submitted and executed, MapReduce is successfully connected to OBS. For example, create a Hive table and set its **Location** to an OBS path.

1. Log in to the node where the Hive client is installed and run the following commands. The authenticated user must have the permission to create Hive tables.

   Go to the client installation directory.

   ```
   cd Client installation directory
   ```

   Load the environment variables.

   ```
   source bigdata_env
   ```

   Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

   ```
   kinit Service user
   ```

2. Log in to the Hive client.

   ```
   beeline
   ```

3. Create a Hive table and insert data into the table.

   a. Create a table.

   ```
   create table test(name string) location "obs://OBS parallel file system name/user/hive/warehouse/Database name/Table name";
   ```

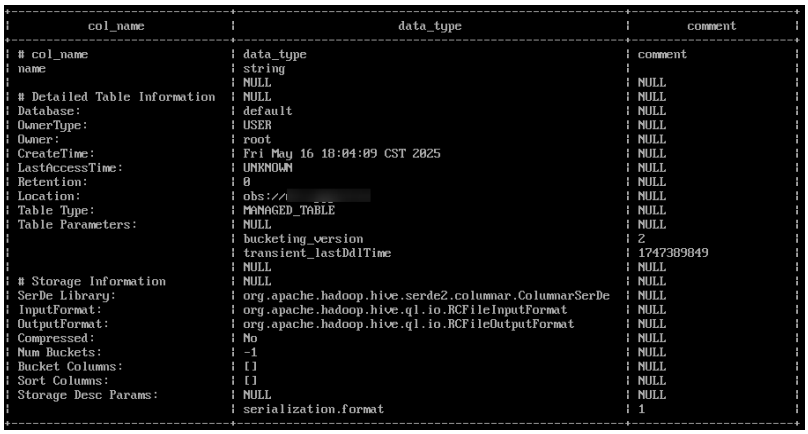
   b. Check the table location.

   ```
   desc formatted test;
   ```

   It is found that the **Location** of the table is the OBS path.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 10-17 Checking the table location



col_name	data_type	comment
# col_name	data_type	comment
name	string	
	NULL	NULL
# Detailed Table Information	NULL	NULL
Database:	default	NULL
OwnerType:	USER	NULL
Owner:	root	NULL
CreateTime:	Fri May 16 18:04:09 CST 2025	NULL
LastAccessTime:	UNKNOWN	NULL
Retention:	0	NULL
Location:	obs://[redacted]	NULL
Table Type:	MANAGED_TABLE	NULL
Table Parameters:	NULL	NULL
	bucketing_version	2
	transient_lastDdlTime	1747389849
	NULL	NULL
# Storage Information	NULL	NULL
SerDe Library:	org.apache.hadoop.hive.serde2.columnar.ColumnarSerDe	NULL
InputFormat:	org.apache.hadoop.hive.ql.io.RCFileInputFormat	NULL
OutputFormat:	org.apache.hadoop.hive.ql.io.RCFileOutputFormat	NULL
Compressed:	No	NULL
Num Buckets:	-1	NULL
Bucket Columns:	[]	NULL
Sort Columns:	[]	NULL
Storage Desc Params:	NULL	NULL
	serialization.format	1

c. Insert data.

```
insert into table test values("A");
```

d. Query table data.

```
select * from test;
```

4. Log in to FusionInsight Manager, choose **Cluster** > **Services** > **Yarn**, and click the hyperlink next to **ResourceManager WebUI** to go to the YARN web UI. On the **All Applications** page, click the application ID of the target Hive task to view the job running information. If **Application Type** is **MAPREDUCE**, the job submitted to OBS is a MapReduce job.
5. Click **Logs** in the **Logs** column at the bottom of the page to view job logs.

----End

10.3.3.7 Interconnecting Presto with OBS Using an IAM Agency

After configuring decoupled storage and compute for a cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can use **presto_cli.sh** client to create tables and store them in OBS.

Interconnecting Presto with OBS (Kerberos authentication is disabled for the cluster (in normal mode))

Step 1 Log in to the node where the client is installed as the client installation user.

For details about how to download and install the cluster client, see .

Step 2 Configure environment variables.

Go to the client installation directory.

```
cd Client installation directory
```

Configure environment variables.

```
source bigdata_env
```

Step 3 Connect to the Presto client.

```
presto_cli.sh
```

Step 4 Create a schema and set **location** to an OBS path.

```
CREATE SCHEMA hive.demo WITH (location = 'obs://mrs-word001/presto-demo002/');
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 5 Create a table in the schema. The table data is stored in the OBS file system.

```
CREATE TABLE hive.demo.demo_table WITH (format = 'ORC') AS SELECT * FROM tpch.sf1.customer;
```

Figure 10-18 Return result

```
[root@node-master2mdc0 ~]# presto_cli.sh
--server http://192.168.3.66:7520
presto> CREATE SCHEMA hive.demo WITH (location = 'obs://mrs-word001/presto-demo/');
CREATE SCHEMA
presto> CREATE TABLE hive.demo.demo_table WITH (format = 'ORC') AS SELECT * FROM tpch.sf1.customer;
CREATE TABLE: 150000 rows

Query 20191221_033019_00001_ukfbz, FINISHED, 2 nodes
Splits: 42 total, 42 done (100.00%)
0:09 [150K rows, 0B] [16K rows/s, 0B/s]
```

Step 6 Exit the client.

```
exit
```

----End

Interconnecting Presto with OBS (Kerberos authentication is enabled for the cluster (in security mode))

Step 1 Log in to FusionInsight Manager of the MRS cluster.

For details about how to log in to FusionInsight Manager, see **Accessing MRS Manager**.

Step 2 Create a role with the Hive Admin Privilege permission, for example, **prestorole**. For how to create a role, see **Managing MRS Cluster Roles**.

Step 3 Create a user that belongs to the Presto and Hive groups and bind the role created in **Step 2** to the user, for example, **presto001**. For how to create a user, see **Creating an MRS Cluster User**.

If a human-machine user is created, log in to MRS Manager as the new user and change the initial password.

Step 4 Download the user credential.

- For versions earlier than MRS 3.*x*, log in to MRS Manager, choose **System** > **Manage User**. Locate the row containing the new user, click **More**, and select **Download Authentication Credential**.
- Log in to FusionInsight Manager for MRS 3.*x* or later, choose **System** > **Permission** > **User**. Locate the row containing the new user, click **More**, and select **Download Authentication Credential**.

Step 5 Decompress the downloaded user credential file and save the obtained **krb5.conf** and **user.keytab** files to the client directory, for example, *Client installation directory*/**Presto/**.

For details about how to download and install the cluster client, see .

Step 6 Log in to the node where the client is installed as the client installation user.

Step 7 Obtain the user principal.

```
klist -kt Client installation directory/Presto/user.keytab
```

Step 8 Configure environment variables.

Go to the client installation directory.

```
cd Client installation directory
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configure environment variables.

```
source bigdata_env
```

Step 9 Connect to the Presto server of the cluster.

```
presto_cli.sh --krb5-config-path {krb5.conf file path} --krb5-principal {User principal} --krb5-keytab-path {user.keytab file path} --user {Presto username}
```

Table 10-7 How to obtain parameter values

Parameter Value	How to Obtain
krb5.conf file path	Replace it with the file path set in **Step 5**, for example, *Client installation directory*/**Presto/krb5.conf**.
user.keytab file path	Replace it with the file path set in **Step 5**, for example, *Client installation directory*/**Presto/user.keytab/opt/Bigdata/client/Presto/user.keytab**.
User principal	Replace it with the result returned in **Step 7**.
Presto username	Replace it with the username created in **Step 3**, for example, **presto001**.

Step 10 Create a schema and set **location** to an OBS path.

```
CREATE SCHEMA hive.demo01 WITH (location = 'obs://mrs-word001/presto-demo002/');
```

Step 11 Create a table in the schema. The table data is stored in the OBS file system.

```
CREATE TABLE hive.demo01.demo_table WITH (format = 'ORC') AS SELECT * FROM tpch.sf1.customer;
```

Figure 10-19 Returned result



Step 12 Exit the client.

```
exit
```

----End

10.3.3.8 Interconnecting Spark with OBS Using an IAM Agency

After configuring decoupled storage and compute for a cluster by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can create tables with OBS paths as their location on the Spark client.

Interconnecting with OBS Using Spark Beeline

Step 1 Log in to the MRS Manager and choose **Cluster** > **Services** > **Spark2x** > **Configurations** > **All Configurations**. For how to log in to MRS Manager, see **Accessing MRS Manager**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

In the left navigation pane, choose **JDBCServer2x** > **Customization**. Add the configuration item **dfs.namenode.acls.enabled** to the **spark.hdfs-site.customized.configs** parameter and set the value to **false** to disable the HDFS ACL function.

Step 2 Search for the **spark.sql.statistics.fallBackToHdfs** parameter and set its value to **false**.

Step 3 Save the configurations and restart the JDBCServer2x instance.

Step 4 Log in to the client installation node as the client installation user.

For details about how to download and install the cluster client, see .

Step 5 Run the following commands to configure environment variables:

```
source Client installation directory/bigdata_env
```

Step 6 Authenticate the user of the cluster with Kerberos authentication enabled. Skip this step for the user of the cluster with Kerberos authentication disabled.

```
kinit Username
```

Step 7 Access OBS using Spark beeline. The following example creates a table named **test** in the **obs://mrs-word001/table/** directory.

Access **spark-beeline**.

```
spark-beeline
```

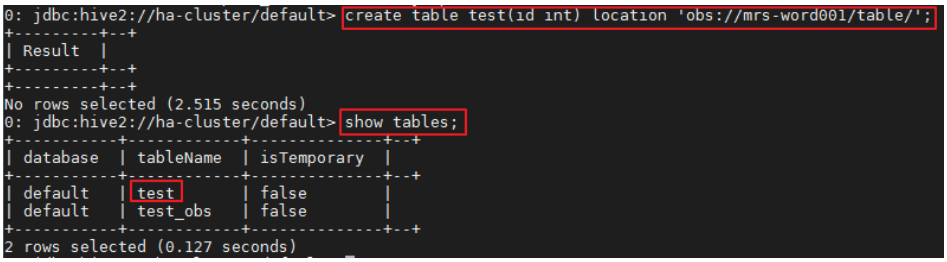
Create the **test** table.

```
create table test(id int) location 'obs://mrs-word001/table/';
```

Step 8 Run the following command to query all tables. If table **test** is returned, OBS access is successful.

```
show tables;
```

Figure 10-20 Returned table names



Step 9 Press **Ctrl+C** to exit Spark beeline.

----End

Interconnecting with OBS Using Spark SQL

Step 1 Log in to the client installation node as the client installation user.

For details about how to download and install the cluster client, see .

Step 2 Run the following commands to configure environment variables:

```
source Client installation directory/bigdata_env
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 3 Modify the configuration file:

```
vim Client installation directory/Spark2x/spark/conf/hdfs-site.xml
```

Modify the following content, where the **dfs.namenode.acls.enabled** parameter specifies whether to enable the HDFS ACL function.

```
<property>
<name>dfs.namenode.acls.enabled</name>
<value>false</value>
</property>
```

Step 4 Authenticate the user of the cluster with Kerberos authentication enabled. Skip this step for the user of the cluster with Kerberos authentication disabled.

```
kinit Username
```

Step 5 Access OBS using Spark SQL CLI. For example, create a table named **test** in the **obs://mrs-word001/table/** directory.

1. Go to the Spark **bin** directory.
   ```
   cd Client installation directory/Spark2x/spark/bin
   ```
2. Log in to the Spark SQL CLI.
   ```
   ./spark-sql
   ```
3. Run the following command in the Spark SQL CLI to create a table:
   ```
   create table test(id int) location 'obs://mrs-word001/table/';
   ```

Step 6 Check whether the table exists.

```
show tables;
```

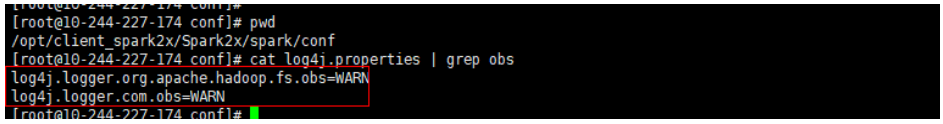
Step 7 Run **exit;** to exit the Spark SQL CLI.

NOTE

If a large number of logs are printed in the OBS file system, the read and write performance may be affected. You can adjust the log level of the OBS client as follows:

1. Go to the **conf** directory.
   ```
   cd Client installation directory/Spark2x/spark/conf
   ```
2. Edit the file **log4j.properties**.
   ```
   vi log4j.properties
   ```
3. Add the OBS log level configuration to the file.
   ```
   log4j.logger.org.apache.hadoop.fs.obs=WARN
   log4j.logger.com.obs=WARN
   ```

Figure 10-21 Adding an OBS log level



----End

Using Spark Shell to Read OBS Files

Step 1 Log in to the client installation node as the client installation user.

For details about how to download and install the cluster client, see .

Step 2 Run the following commands to configure environment variables:

```
source Client installation directory/bigdata_env
```

Step 3 Modify the configuration file:

```
vim Client installation directory/Spark2x/spark/conf/hdfs-site.xml
```

Modify the following content, where the **dfs.namenode.acls.enabled** parameter specifies whether to enable the HDFS ACL function.

```
<property>
<name>dfs.namenode.acls.enabled</name>
<value>false</value>
</property>
```

Step 4 Authenticate the user of the cluster with Kerberos authentication enabled. Skip this step for the user of the cluster with Kerberos authentication disabled.

```
kinit Username
```

Step 5 Run the following commands to log in to the Spark SQL CLI:

Go to the **bin** directory.

```
cd Client installation directory/Spark2x/spark/bin
```

Log in to the Spark SQL CLI.

```
./spark-sql
```

Step 6 Create a table in OBS and import data to the table.

1. Create a database.
   ```
   create database test location "obs://Parallel file system path/test";
   ```
2. Switch to the new database.
   ```
   use test;
   ```
3. Create a table.
   ```
   create table test1(a int,b int) using parquet;
   ```
4. Insert data into the table.
   ```
   insert into test1 values(1,2);
   ```
5. View table details.
   ```
   desc formatted test1;
   ```

   The **Location** of the table shown in **Figure 10-22** is an OBS path.

   Figure 10-22 Checking the location of the table

   ```
   spark-sql> desc formatted test11;
   a       int     NULL
   b       int     NULL

   # Detailed Table Information
   Database        test1
   Table   test11
   Owner   root
   Created Time    Tue Nov 21 10:35:40 CST 2023
   Last Access     UNKNOWN
   Created By      Spark                -315008
   Type    MANAGED
   Provider        parquet
   Location        obs://        /test11/test11
   Serde Library   org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe
   InputFormat     org.apache.hadoop.hive.ql.io.parquet.MapredParquetInputFormat
   OutputFormat    org.apache.hadoop.hive.ql.io.parquet.MapredParquetOutputFormat
   Time taken: 0.235 seconds, Fetched 16 row(s)
   spark-sql>
   ```

6. Run **exit;** to exit the Spark SQL CLI.

Step 7 Run the following command to go to the Spark **bin** directory:

```
cd Client installation directory/Spark2x/spark/bin
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 8 Log in to the Spark SQL CLI.

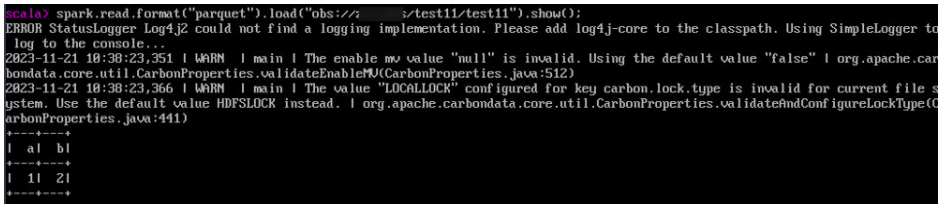
```
./spark-shell
```

Step 9 In the Spark Shell CLI, run the following command to query the table created in **Step 5.3**:

```
spark.read.format("parquet").load("obs://Parallel file system path/test1").show();
```

The table data is displayed, as shown below.

Figure 10-23 Viewing table data



Step 10 Run the **:quit** command to exit the Spark Shell CLI.

----End

10.3.3.9 Interconnecting Sqoop with OBS Using an IAM Agency

After connecting the Sqoop client to an OBS file system by referring to **Interconnecting an MRS Cluster with OBS Using an IAM Agency**, you can import tables from a relational database to OBS or export tables from OBS to a relational database on the Sqoop client.

Prerequisites

The RDS for MySQL instance and the MRS cluster where Sqoop is located are in the same VPC.

Importing MySQL Data to OBS Using Sqoop

Step 1 Log in to the RDS console, choose **Instances**, locate the row that contains the target RDS for MySQL instance, click **Log In** in the **Operation** column to log in to the instance as user **root**.

Step 2 On the homepage, click **Create Database** to create a database, for example, **test**. In the database list, click the name of the newly created database to go to the database management page. On the **Tables** page, click **Create Table** to create a table, for example, **sourcetable**, which contains the **id** and **name** fields.

Step 3 Click **Query SQL Statements** in the row where the new table is located. On the displayed page, switch to the database where the table is located and run the following SQL command to insert data into the table:

```
INSERT INTO sourcetable (id,name) VALUES (11,"A");
INSERT INTO sourcetable (id,name) VALUES (22,"B");
```

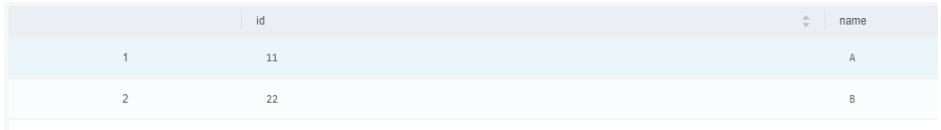
Query table data.

```
select * from `sourcetable`;
```

The queried table data is as follows:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

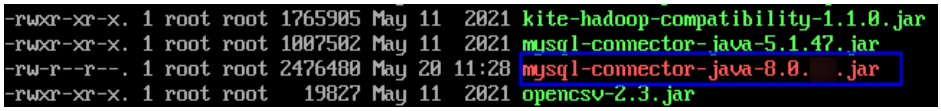
Figure 10-24 Querying table data



	id	name
1	11	A
2	22	B

Step 4 Download the MySQL driver package of the required version from the MySQL official website **https://downloads.mysql.com/archives/c-j/**, decompress the package, and upload it to the *Client installation directory*/**Sqoop/sqoop/lib** directory on the node where the Sqoop client is installed.

Figure 10-25 Uploading the MySQL driver file



Step 5 Log in to the node where the client is installed as the client installation user.

For details about how to download and install the cluster client, see .

Step 6 Configure environment variables and authenticate the user.

Go to the client installation directory.

```
cd Client installation directory
```

Load the environment variables.

```
source bigdata_env
```

Authenticate the user. Skip this step for clusters with Kerberos authentication disabled.

```
kinit Component service user
```

Step 7 Import MySQL table data to OBS using **sqoop import**.

```
sqoop import --connect jdbc:mysql://MySQL IP address:3306/test --username root --password xxx --table sourcetable --target-dir obs://Parallel file system name/xxx --delete-target-dir --fields-terminated-by "," -m 1 --as-textfile
```

Step 8 Log in to OBS console, choose **Parallel File System**, go to the parallel file system path for storing MySQL data configured in **Step 7**, and download and check whether the data file is the same as that inserted into the MySQL table in **Step 3**.

```
11,A
22,B
```

----End

10.3.4 Configuring Fine- Grained OBS Access Permissions for MRS Cluster Users

When fine-grained permission control is enabled, you can configure OBS access permissions to implement access control on directories in OBS file systems.

This function enables you to control MRS users' access to OBS resources. For example, if you allow user group A to only access log files in a specified OBS file system, perform the following operations:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Configure an agency with OBS access permissions for an MRS cluster so that OBS can be accessed using the temporary AK/SK automatically obtained by the ECS.
2. Create a policy on the IAM console to allow access to log files in a specified OBS file system, and create an agency bound to the policy permission.
3. In the MRS cluster, bind the new agency to user group A so that user group A only has the permission to access log files in the specified OBS file system.

In the following scenarios, the username used for submitting jobs is an internal username so that MRS multi-user access to OBS is not supported.

- In a cluster with Kerberos authentication enabled, the built-in username of **spark-beeline** for submitting jobs is **spark**. In a cluster with Kerberos authentication disabled, the built-in username is **omm**.
- In a cluster with Kerberos authentication enabled, the built-in usernames of Presto for submitting jobs is **omm** and **hive**. In a cluster with Kerberos authentication disabled, the built-in username is **omm**. On the cluster details page of the MRS console, choose **Components** > **Presto** > **Service Configuration**, set **Type** to **All**, search for **hive.hdfs.impersonation.enabled**, and change its value to **true**. In this way, multiple MRS users can access OBS with fine-grained permissions.

Notes and Constraints

This section does not apply to MRS 1.9.2.

Prerequisites

Fine-grained permission control has been enabled. For details about permissions management, see **Creating an IAM User and Granting MRS Permissions**.

Configuring an Agency with OBS Access Permissions for a Cluster

Follow instructions in **Interconnecting an MRS Cluster with OBS Using an IAM Agency** to configure an agency with OBS access permissions.

The agency takes effect for all users (including internal users) and user groups in the cluster. To control the permissions of users and user groups in the cluster to access OBS, perform the following operations.

When configuring permissions on an OBS path, if you grant write permissions, you must also configure the corresponding recycle bin path. The default recycle bin path is **/user/***${current.user}***/.Trash/**, where *${current.user}* indicates the current username.

Creating a Policy and an Agency on IAM

Create policies with different access permissions and bind the policies to the agency. For details, see **Reference: Creating a Policy and an Agency on IAM**.

Configuring OBS Permission Control Mapping

Step 1 On the MRS management console, choose **Active Clusters** and click the cluster name.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 2 In the **Basic Information** area on the **Dashboard** tab page, click **Manage** next to **OBS Permission Control**.

Step 3 Click **Add Mapping** and set parameters according to **Table 10-8**.

Table 10-8 Adding an OBS permission control mapping

Parameter	Description
IAM Agency	Select the agency created in **Step 2**.
Type	• **User**: User-level mapping • **Group**: User group-level mapping User-level mapping takes priority over user group-level mapping. If you select **Group**, you are advised to enter the primary group name in **MRS User (User Group)**. Do not use the same username (group) in multiple mapping records.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
MRS User (User Group)	Use commas (,) to separate multiple names of users or user groups. • If OBS permission control is not configured for a user and no AK/SK is set, the OBS Operator permission in **MRS_ECS_DEFAULT_AGENCY** will be used to access OBS. It is recommended that you do not bind the internal user of a component to any agency. • If you need to configure an agency for a component's internal user when submitting a job in the following scenarios, observe the following requirements: – To control permissions on spark-beeline operations, configure the username **spark** for clusters with Kerberos authentication enabled and the username **omm** for clusters with Kerberos authentication disabled. – To control permissions on Presto operations, configure usernames **omm**, **hive**, and the one used to log in to the client for clusters with Kerberos authentication enabled, and configure usernames **omm** and the one used to log in to the client for clusters with Kerberos authentication disabled. – If you want to use Hive to create tables in beeline mode, configure the internal user **hive**.

Step 4 Click **OK**.

Step 5 Select **I agree to authorize the trust relationships between MRS Users (Groups) and IAM agencies**, and click **OK**. The mapping between the MRS user and OBS permission is added.

If appears next to **OBS Permission Control** on the **Dashboard** tab page or the mapping table has been updated for OBS permission control, the mapping takes effect. It takes about 1 minute to for the mapping to take effect.

In the **Operation** column of the mapping list, you can edit or delete the added mapping.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- If OBS permission control is not configured for a user and no AK/SK is set, OBS access is granted using the permissions of the agency configured for the cluster in the **Object Storage Service (OBS)** project.
- Once AK/SK credentials are configured, OBS access is performed using those credentials, regardless of whether OBS permission control is configured.
- Security Administrator permissions are required to modify, create, or delete a mapping.
- To apply the mapping changes in Spark Beeline, Hive Beeline, and Presto, you must restart Spark, exit and re-enter Beeline, and restart Presto, respectively.

----End

Component Access to OBS When OBS Permission Control Is Enabled

Step 1 Log in to any node in a cluster as user **root** using the password set during cluster creation.

Step 2 Run the following commands to set the environment variables:

cd *Client installation directory*

source *Client installation directory*/**bigdata_env**

Step 3 If the Kerberos authentication is enabled for the current cluster, run the following command to authenticate the user. If the Kerberos authentication is disabled for the current cluster, skip this step:

kinit *MRS cluster user*

Example:

kinit admin

Step 4 If Kerberos authentication is disabled for the current cluster, run the following commands to log in as a user who belongs to the **supergroup** group. Replace *XXXX* with the username. For how to create a user, refer to **Creating an MRS Cluster User**.

mkdir /home/*XXXX*

chown *XXXX* **/home/***XXXX*

su - *XXXX*

Step 5 Access OBS. You do not need to configure AK/SK credentials or endpoints. The OBS path format is **obs://***OBS parallel file system name*/*XXX*.

hadoop fs -ls "obs://obs-example/job/hadoop-mapreduce-examples-3.1.2.jar**"**

Note that:

- To delete files from OBS using the **hadoop fs** command, run **hadoop fs -rm -skipTrash**.
- If table creation in Spark SQL or Spark Beeline does not involve data import, OBS access is not triggered. That is, if you create a table in an OBS directory on which you do not have permission, the **CREATE TABLE** operation will still

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

be successful, but the error message "**403 AccessDeniedException**" is displayed when you insert data.

----End

Reference: Creating a Policy and an Agency on IAM

Step 1 Create a policy on IAM.

1. Log in to the IAM console.
2. In the navigation pane on the left, choose **Permissions** > **Policies/Roles**. On the displayed page, click **Create Custom Policy**.
3. Set parameters according to **Table 10-9**.

Table 10-9 Policy parameters

Parameter	Description
Policy Name	Only letters, digits, spaces, and special characters (-_.,) are allowed.
Scope	Select **Global services**, because OBS is a global service.
Policy View	Select **Visual editor**.
Policy Content	– **Allow**: Select **Allow**. – **Select service**: Select **Object Storage Service (OBS)**. – **Select action**: Select **ReadWrite**, **ReadOnly**, and **ListOnly**. – **Resources** under **All**: Select **Specific** and set the following parameters: ▪ **object**: Select **Specify resource path** and click **Add Resource Path** to add a path. The **/tmp** directory is used as an example. For example, enter *obs_bucket_name***/tmp/** and *obs_bucket_name***/tmp/***. ▪ **bucket**: Select **Specify resource path**, click **Add Resource Path**, and enter *obs_bucket_name*. – (Optional) Request condition: not added.
Description	(Optional) Brief description about the policy.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the data write operation of each component is implemented in **rename** mode, the permission to delete objects must be configured when data is written.

4. Click **OK** to save the policy.

Step 2 Create an agency on IAM.

1. Log in to the IAM console.
2. Choose **Agencies**. On the displayed page, click **Create Agency**.
3. Set parameters according to **Table 10-10**.

Table 10-10 Agency parameters

Parameter	Description
Agency Name	Only letters, digits, spaces, and special characters (-_.,) are allowed.
Agency Type	Select **Common account**.
Delegated Account	Enter your cloud account, that is, the account you register using your mobile phone number. It cannot be a federated user or an IAM user created using your cloud account.
Validity Period	Set this parameter as required.
Description	(Optional) Brief description about the agency.
Permissions	1. In the **Project [Region]** column, locate the row where **OBS** is, click **Attach Policy**. 2. Select the policy created in **Step 1** to display it in **Selected Policies**. 3. Click **OK**.

4. Click **OK** to save the agency.

   If you modify an agency or its bound policies after it has been used to access OBS, the changes take effect within 15 minutes.

----End

10.3.5 How Do I Access OBS Using an MRS Client Installed Outside a Cluster?

Scenarios

In storage-compute decoupling scenarios where data is computed in an MRS cluster and stored in OBS buckets, you can obtain a temporary AK/SK using an agency and then use the AK/SK to access the OBS server. To access OBS from a

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

client on a node outside the cluster, obtain an AK/SK through Guardian. Guardian is a component developed by the MRS team that allows clients outside the cluster to access OBS using temporary AKs/SKs.

Notes and Constraints

This function is available in MRS 3.1.5 or later.

Accessing OBS Through Guardian on a Client Installed Outside the Cluster

Step 1 Install Guardian in the cluster.

- For a new cluster in creation, select **Guardian**.
- For an existing cluster, add the Guardian component in the **Components** tab.
  a. On the cluster details page, choose **Components** and click **Add Service**.
  b. In the service list, select the services to be added and click **Next**.
  c. On the **Topology Adjustment** page, select the node where the service is to be deployed. (You are advised to deploy the Guardian service on the Master node.)
  d. Click **OK**. After the service is added, you can view the added service on the **Components** page.

     The services added on the MRS console are automatically synchronized to MRS Manager.

Step 2 Configure storage and compute decoupling for the cluster by referring to **Configuring a Storage-Compute Decoupled Cluster (Agency)** or **Configuring Fine-Grained Permissions for MRS Multi-User Access to OBS**.

Step 3 Install or update the client.

- For details about how to install the client on a node outside the cluster, see **Installing a Client (MRS 3.x)**.
- For details about how to update an existing client, see **Updating the MRS Cluster Client After the Server Configuration Expires**.

Step 4 Once the installation is successful, you can access OBS. For example, access OBS from the HDFS client of an MRS cluster with Kerberos authentication disabled:

1. Log in to the node where the client is installed as the client installation user.
2. Go to the client installation directory and configure the environment variables:

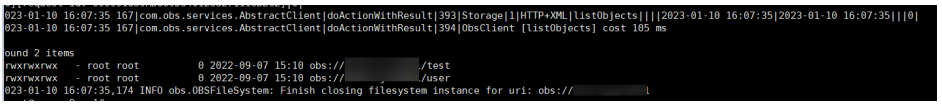
   cd *Client installation directory*

   source bigdata_env
3. Run the following command on the HDFS client to access OBS:

   hdfs dfs -ls obs://*Directory of the OBS parallel file system*

   If information similar to the following is displayed, the interconnection is successful:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11 FAQs

11.1 MRS Basics

11.1.1 What Is MRS Used For?

Question:

What Can MapReduce Service (MRS) Do?

Answer:

MapReduce Service (MRS) is an enterprise-grade big data platform that allows you to quickly build and operate economical, secure, full-stack, cloud-native big data environments on the cloud. It provides engines such as ClickHouse, Spark, Flink, Kafka, and HBase, and supports convergence of data lake, data warehouse, business intelligence (BI), and artificial intelligence (AI). Fully compatible with open-source components, MRS helps you rapidly innovate and expand service growth.

11.1.2 What Types of Distributed Storage Does MRS Support?

Question:

What Types of Distributed Storage Do MRS Clusters Support? What Versions Are Available?

Answer:

The MRS clusters use the mainstream big data component Hadoop. Currently, Hadoop 3.x is supported and will be updated as the cluster evolves.

MRS allows you to store data in OBS and use an MRS cluster for data computing only. In this way, storage and compute are decoupled.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.1.3 What Are Regions and AZs?

A region and an availability zone (AZ) identify the location of a data center. You can create resources in a specific AZ of a region.

- A region is a physical data center. Each region is completely isolated to ensure high fault tolerance and stability. After creating resources in a region, you cannot change the region.
- An AZ is a physical location with independent power supplies and network in a region. A region contains one or more AZs that are physically isolated but interconnected through the internal network. The fault of an AZ will not affect other AZs. The internal network provides economical connection with low latency.

What Are the Considerations for Selecting a Region?

When selecting a region, consider the following factors:

- Location

  You are advised to select a region closest to your target users. This reduces network latency and improves access rate. However, there is no distinct difference between the infrastructure, BGP network quality, and resource operations and configurations in the same region/country. Therefore, you do not need to consider the network latency when selecting a region in a country.
- Resource price

  Resource prices may vary in different regions.

What Are the Considerations for Selecting an AZ?

Consider the requirements for disaster recovery (DR) and network latency when selecting an AZ:

- Deploy resources in different AZs in the same region for DR purposes.
- Deploy resources in the same AZ for minimum latency.

11.1.4 Can I Change the Network Segment of Nodes in an MRS Cluster?

Question:

Can I Change the Network Segment of Nodes in an MRS Cluster?

Answer:

Yes, you can. Perform the following steps to change the network segment of nodes in a cluster:

1. On the **Dashboard** page of the cluster, locate the **Network Information** pane, and click **Change Subnet** on the right of subnet name.
2. Select another subnet in the VPC to which the current cluster belongs to expand the available subnet IP addresses.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Selecting a new subnet will not change the IP addresses and subnets of existing nodes.

11.1.5 Can I Downgrade the Specifications of an MRS Cluster Node?

Question:

Can I Downgrade the Specifications of an MRS Cluster Node?

Answer:

Currently, the configuration specifications of nodes in an MRS cluster cannot be downgraded.

11.1.6 Are Hive Components of Different Versions Compatible with Each Other?

Hive 3.1 is incompatible with Hive 1.2 in the following items:

- Strings cannot be converted to int in Hive 3.1.
- The user-defined functions (UDFs) of the **Date** type are changed to Hive 3.1 built-in UDFs.
- Hive 3.1 does not provide the index function anymore.
- The JDBC drivers in Hive 3.1 and Hive 1.2 are incompatible.
- In Hive 3.1, column names in ORC files are case-sensitive and underscores-sensitive.
- Hive 3.1 does not allow columns named **time**.

11.1.7 What Are the Differences Between OBS and HDFS in Data Storage?

The data processed by MRS is from OBS or HDFS. OBS is an object-based storage service that provides secure, reliable, and cost-effective storage of huge amounts of data. MRS can directly process data in OBS. You can view, manage, and use data by using the OBS console or OBS client. In addition, you can use REST APIs independently or integrate APIs to service applications to manage and access data.

- Data stored in OBS: Data storage is decoupled from compute. The cluster storage cost is low, and storage capacity is not limited. Clusters can be deleted at any time. However, the computing performance depends on the OBS access performance and is lower than that of HDFS. OBS is recommended for applications that do not demand a lot of computation.
- Data stored in HDFS: Data storage is not decoupled from compute. The cluster storage cost is high, and storage capacity is limited. The computing performance is high. You must export data before you delete clusters. HDFS is recommended for computing-intensive scenarios.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.1.8 What Are the Solutions for Processing over 1 Billion Data Records?

Question:

What Are the Solutions for Processing over 1 Billion Data Records?

Answer:

- You are advised to use TaurusDB databases in scenarios where data is updated, online transaction processing (OLTP) is performed, and complex analysis is required.
- Impala and Kudu in MRS also meet this requirement. Impala and Kudu can load all join tables to the memory in the join operation.

11.1.9 What Are the Advantages of ZSTD?

Question:

What are the advantages of ZSTD?

Answer:

Zstandard (ZSTD) is an open-source lossless data compression algorithm. Its compression performance and compression ratio are better than those of other compression formats supported by Hadoop.

For details, see **https://github.com/L-Angel/compress-demo**.

11.2 Cluster Creation

11.2.1 How Do I Create an MRS Cluster Using a Custom Security Group?

Question:

How Do I Create an MRS Cluster Using a Custom Security Group?

Answer:

You can create an MRS cluster using a user-defined security group in either of the following ways:

- If you want to use a security group created by yourself, you need to enable port 9022 when you create a cluster.
- When you create a cluster, set **Security Group** to **Auto create**.

11.2.2 What Should I Do If HDFS, Yarn, and MapReduce Components Are Unavailable When I Create an MRS Cluster?

Question:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

What Should I Do If HDFS, Yarn, and MapReduce Components Are Unavailable When I Create an MRS Cluster?

Answer:

The HDFS, YARN, and MapReduce components are integrated in Hadoop. If the three components are unavailable when are MRS cluster is created, select Hadoop instead. After an MRS cluster is created, HDFS, YARN, and MapReduce are available in the **Components** page.

11.2.3 What Should I Do If the ZooKeeper Component Is Unavailable When I Create an MRS Cluster?

Question:

Why Cannot I Find the ZooKeeper Component When I Create an MRS Cluster?

Answer:

- If you create a cluster of a version earlier than MRS 3.*x*, ZooKeeper is installed by default and is not displayed on the GUI.

  After the cluster is created, the ZooKeeper component is available on the **Components** page.
- If you create a cluster of MRS 3.*x* or later, ZooKeeper is available on the GUI and is selected by default.

11.3 Web Page Access

11.3.1 How Do I Change the Session Timeout Duration for an Open Source Component Web UI?

You can perform the following operations to configure the timeout duration of the MRS Manager web UI and an open-source component web UI.

CAUTION

Set the minimum session duration based on service requirements. Otherwise, there will be security risks.

Changing the Timeout Duration of the MRS Manager Web UI and Authentication Center Page

- **Clusters of MRS 3.*x* (MRS 3.3.0 or later)**:
  a. Log in to MRS Manager and choose **System** > **OMS**.
  b. In the list, locate the row that contains **tomcat** and click **Modify Configuration**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

c. On the displayed page, set **Session Timeout** as required and click **OK**. Verify the user password and wait until the configuration to complete. Once done, the timeout setting will take effect for Manager.

- **Clusters of MRS 3.*x* (Versions earlier than MRS 3.3.0)**:

  a. Log in to the active and standby OMS nodes of the cluster as user **omm** and perform steps 2 to 5.

  b. In the files **/opt/Bigdata/om-server/tomcat/webapps/web/WEB-INF/web.xml** and **/opt/Bigdata/om-server/tomcat/conf/web.xml**, locate the following content, where **20** indicates the session timeout duration, in minutes. You can change the value as required, but it cannot exceed 480.

  ```
  <session-timeout>20</session-timeout>
  ```

  c. In the file **/opt/Bigdata/om-server/tomcat/webapps/cas/WEB-INF/classes/config/application.properties**, add the following information, where **28800** indicates the validity period of the authentication center, in seconds. You can change the value as required, but it cannot exceed 28800.

  ```
  ticket.tgt.timeToKillInSeconds=28800
  ```

  d. In the file **/opt/Bigdata/om-server/tomcat/webapps/cas/WEB-INF/classes/apereo.properties**, add the following content, where **28800** indicates Ticket Granting Ticket (TGT) validity period, in seconds. You can change the value as required, but it cannot exceed 28800.

  ```
  cas.ticket.tgt.maxTimeToLiveInSeconds=28800
  cas.ticket.tgt.timeToKillInSeconds=28800
  ```

  e. In the file **/opt/Bigdata/om-server/tomcat/webapps/cas/WEB-INF/classes/bootstrap.properties**, add the following content, where **28800** indicates Ticket Granting Ticket (TGT) validity period, in seconds. You can change the value as required, but it cannot exceed 28800.

  ```
  cas.ticket.tgt.maxTimeToLiveInSeconds=28800
  cas.ticket.tgt.timeToKillInSeconds=28800
  ```

  f. Run the following command on the active management node as user **omm** to restart the Tomcat node:

  sh ${BIGDATA_HOME}/om-server/tomcat/bin/shutdown.sh;sh ${BIGDATA_HOME}/om-server/tomcat/bin/startup.sh

- **For clusters of versions earlier than MRS 3.*x*:**

  a. Log in to the active and standby OMS nodes of the cluster as user **omm** and perform steps 2 to 4.

  b. In the file **/opt/Bigdata/apache-tomcat-xxx/webapps/cas/WEB-INF/web.xml**, locate the following content, where the value **20** indicates the session timeout duration, in minutes. You can change the value as required, but it cannot exceed 480.

  ```
  <session-timeout>20</session-timeout>
  ```

  c. In the file **/opt/Bigdata/apache-tomcat-xxx/webapps/web/WEB-INF/web.xml**, locate the following content, where the value **20** indicates the session timeout duration, in minutes. You can change the value as required, but it cannot exceed 480.

  ```
  <session-timeout>20</session-timeout>
  ```

  d. In the file **/opt/Bigdata/apache-tomcat-xxx/webapps/cas/WEB-INF/spring-configuration/ticketExpirationPolicies.xml**, add the following content, where the value **1200** indicates the validity period of the authentication center, in seconds. You can change the value as required, but it cannot exceed 28800.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
p:maxTimeToLiveInSeconds="${tgt.maxTimeToLiveInSeconds:1200}"
p:timeToKillInSeconds="${tgt.timeToKillInSeconds:1200}"
```

e. Restart the Tomcat node on the active master node.

i. On the active master node, run the **netstat -anp |grep 28443 |grep LISTEN | awk '{print $7}'** command as user **omm** to query the Tomcat process ID.

ii. Run the **kill -9 {*pid*}** command, where *{pid}* indicates the Tomcat process ID obtained in the previous step.

iii. Wait for the process to restart. You can run the **netstat -anp |grep 28443 |grep LISTEN** command to check whether the process is restarted. If the process can be queried, the process is successfully restarted. If the process cannot be queried, query the process again later.

Changing the Web UI Timeout Duration of a Component

- **For clusters of versions earlier than MRS 3.*x*:**

a. On the cluster details page, choose **Components** > **meta** > **Configurations**.

b. Click **All Configurations** tab and search for **http.server.session.timeout.secs** on the displayed page. If the parameter exists, change its value as required. If not, the version does not support dynamic adjustment of the session duration.

c. Restart the meta component for the modification to take effect.

11.3.2 What Can I Do If the Dynamic Resource Plan Page in MRS Tenant Management Cannot Be Refreshed?

Question:

What Can I Do If the Dynamic Resource Plan Page in MRS Tenant Management Cannot Be Refreshed?

Answer:

Step 1 Log in to the Master1 and Master2 nodes as user **root**.

Step 2 Run the **ps -ef |grep aos** command to check the AOS process ID.

Step 3 Run the **kill -9** *AOS process ID* command to end the AOS process.

Step 4 Wait until the AOS process is automatically restarted. You can run the **ps -ef |grep aos** command to check whether the AOS process exists. If the process exists, the restart is successful. If the process does not exist, retry later.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.3.3 What Do I Do If the Kafka Topic Monitoring Tab Is Not Displayed on Manager?

Question:

What Do I Do If the Kafka Topic Monitoring Tab Is Not Displayed on Manager?

Answer:

Step 1 Log in to the Master nodes of the cluster and switch to user **omm**.

Step 2 Go to the **/opt/Bigdata/apache-tomcat-7.0.78/webapps/web/WEB-INF/lib/components/Kafka/** directory.

cd /opt/Bigdata/apache-tomcat-*7.0.78***/webapps/web/WEB-INF/lib/components/Kafka/**

Step 3 Copy the ZooKeeper package to the directory.

cp /opt/share/zookeeper-3.5.1-mrs-2.0/zookeeper-3.5.1-mrs-2.0.jar ./

Step 4 Restart Tomcat.

sh /opt/Bigdata/apache-tomcat-*7.0.78***/bin/shutdown.sh**

sh /opt/Bigdata/apache-tomcat-*7.0.78***/bin/startup.sh**

----End

11.3.4 Why Can't I Find the User Management Page on MRS Manager?

Question:

Why can't I find the user management page on MRS Manager?

Answer:

Check whether you have the **Manager_administrator** permission. If you do not have this permission, **Manage User** will not be available on the **System** page of MRS Manager.

Log in to MRS Manager as user **admin** or another user who has the Manager administration permission.

11.3.5 What Should I Do If the Excel File Downloaded by Hue Cannot Be Opened?

NOTE

This section applies only to versions earlier than MRS 3.*x*.

1. Log in to a Master node as user **root** and switch to user **omm**.

   su - omm

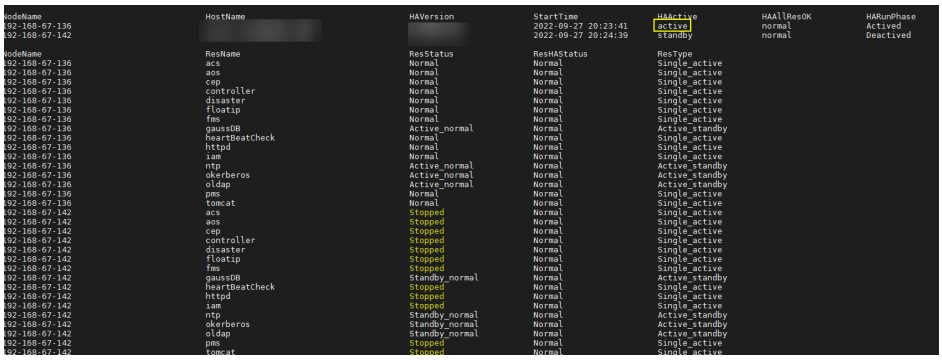
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Check whether the current node is the active OMS node.

   sh ${BIGDATA_HOME}/om-0.0.1/sbin/status-oms.sh

   If **active** is displayed in the command output, the node is the active node. Otherwise, log in to the other Master node.

   Figure 11-1 Active OMS node



3. Go to the **{BIGDATA_HOME}/Apache-httpd-*/conf** directory.

   cd ${BIGDATA_HOME}/Apache-httpd-*/conf

4. Open the **httpd.conf** file.

   vim httpd.conf

5. Search for **21201** in the file and delete the following content from the file. (The values of *proxy_ip* and *proxy_port* are the same as those in the actual environment.)

   ProxyHTMLEnable On
   SetEnv PROXY_PREFIX=https://[proxy_ip]:[proxy_port]
   ProxyHTMLURLMap (https?:\/\/[^:]*:[0-9]*.*) ${PROXY_PREFIX}/proxyRedirect=$1 RV

   Figure 11-2 Content to be deleted



6. Save the modification and exit.
7. Open the **httpd.conf** file again, search for **proxy_hue_port**, and delete the following content:

   ProxyHTMLEnable On
   SetEnv PROXY_PREFIX=https://[proxy_ip]:[proxy_port]
   ProxyHTMLURLMap (https?:\/\/[^:]*:[0-9]*.*) ${PROXY_PREFIX}/proxyRedirect=$1 RV

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 11-3 Content to be deleted



8. Save the modification and exit.
9. Run the following command to restart the **httpd** process:

   sh ${BIGDATA_HOME}/Apache-httpd-*/setup/restarthttpd.sh
10. Check whether the **httpd.conf** file on the standby Master node is modified. If the file is modified, no further action is required. If the file is not modified, modify the **httpd.conf** file on the standby Master node in the same way. You do not need to restart the **httpd** process.
11. Download the Excel file again. You can open the file successfully.

11.4 Authentication and Permission

11.4.1 What Is the User for Logging in to FusionInsight Manager?

Question:

What Is the User for Logging in to FusionInsight Manager?

Answer:

The default account for logging in to Manager is **admin**, and the password is the one you set when you created the cluster.

11.4.2 How Do I Query and Change the Password Validity Period of a User In a Cluster?

Querying the Password Validity Period

Querying the password validity period of a component running user (human-machine user or machine-machine user):

Step 1 Log in to the node where the client is installed as the client installation user.

Step 2 Run the following command to switch to the client directory, for example, **/opt/Bigdata/client**:

cd /opt/Bigdata/client

Step 3 Run the following command to configure environment variables:

source bigdata_env

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 Run the following command and enter the password of user **kadmin/admin** to log in to the kadmin console:

kadmin -p kadmin/admin

NOTE

The default password of user **kadmin/admin** is **Admin@123**. Change the password upon your first login or as prompted and keep the new password secure.

Step 5 Run the following command to view the user information:

getprinc *Internal system username*

Example: **getprinc user1**

```
kadmin:  getprinc user1
......
Expiration date: [never]
Last password change: Sun Oct 09 15:29:54 CST 2022
Password expiration date: [never]
......
```

----End

Querying the password validity period of an OS user:

Step 1 Log in to any master node in the cluster as user **root**.

Step 2 Run the following command to view the password validity period (value of **Password expires**):

chage -l *Username*

For example, to view the password validity period of user **root**, run the **chage -l root** command. The command output is as follows:

```
[root@xxx ~]#chage -l root
Last password change                                    : Sep 12, 2021
Password expires                                        : never
Password inactive                                       : never
Account expires                                         : never
Minimum number of days between password change          : 0
Maximum number of days between password change          : 99999
Number of days of warning before password expires       : 7
```

----End

Changing the Password Validity Period

- The password of a machine-machine user is randomly generated and never expires by default.
- The password validity period of a human-machine user can be changed by modifying the password policy on Manager.

11.4.3 Does an MRS Cluster Support Access Permission Control If Kerberos Authentication Is not Enabled?

For MRS cluster 2.1.0 or earlier, choose **System** > **Configuration** > **Permission** on MRS Manager.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

For MRS cluster 3.*x* or later, choose **System** > **Permission** on FusionInsight Manager.

11.4.4 How Do I Add Tenant Management Permission to Users in a Cluster?

You can assign tenant management permission only in analysis or hybrid clusters, but not in streaming clusters.

The operations vary depending on the MRS cluster version:

Procedure for versions earlier than MRS cluster 3.x:

Step 1 Log in to MRS Manager.

Step 2 Choose **System** > **Manage User**. Select the new user, and click **Modify** in the **Operation** column.

Step 3 In **Assign Rights by Role**, click **Select and Add Role**.

- If you bind the **Manager_tenant** role to the account, the account will have permission to view tenant management information.
- If you bind the **Manager_administrator** role to the account, the account will have permission to view and perform tenant management.

Step 4 Click **OK**.

----End

Procedure for MRS cluster 3.x and later versions:

Step 1 Log in to FusionInsight Manager and choose **System** > **Permission** > **User**.

Step 2 Locate the user and click **Modify**.

Modify the parameters based on service requirements.

If you bind the **Manager_tenant** role to the account, the account will have permission to view tenant management information. If you bind the **Manager_administrator** role to the account, the account will have permission to view and perform tenant management.

NOTE

It takes about three minutes for the settings to take effect after user group or role permission are modified.

Step 3 Click **OK**.

----End

11.4.5 Can I Configure Account Permissions on the Hue Web UI?

Question:

Can I configure account permissions on the Hue web UI in an MRS cluster?

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Answer:

Account permissions cannot be configured on the Hue web UI.

You can configure cluster user roles and user groups on FusionInsight Manager to manage user permissions.

11.4.6 How Do I View the Hive Table Created by Another User?

Versions earlier than MRS 3.*x*:

1. Log in to MRS Manager and choose **System** > **Permission** > **Manage Role**.
2. Click **Create Role**, and set **Role Name** and **Description**.
3. In the **Permission** table, choose **Hive** > **Hive Read Write Privileges**.
4. In the database list, click the name of the database where the table created by user B is stored. The table is displayed.
5. In the **Permission** column of the table created by user B, select **SELECT**.
6. Click **OK**, and return to the **Role** page.
7. Choose **System** > **Manage User**. Locate the row containing user A, click **Modify** to bind the new role to user A, and click **OK**. After about 5 minutes, user A can access the table created by user B.

MRS 3.*x* or later:

1. Log in to FusionInsight Manager and choose **Cluster** > **Services**. On the page that is displayed, choose **Hive**. On the displayed page, choose **More**, and check whether **Enable Ranger** is grayed out.
   - If yes, go to **9**.
   - If no, perform **2** to **8**.
2. Log in to FusionInsight Manager and choose **System** > **Permission** > **Role**.
3. Click **Create Role**, and set **Role Name** and **Description**.
4. In the **Configure Resource Permission** table, choose *Name of the desired cluster* > **Hive** > **Hive Read Write Privileges**.
5. In the database list, click the name of the database where the table created by user B is stored. The table is displayed.
6. In the **Permission** column of the table created by user B, select **Select**.
7. Click **OK**, and return to the **Role** page.
8. Choose **Permission** > **User**. On the **Local User** page that is displayed, locate the row containing user A, click **Modify** in the **Operation** column to bind the new role to user A, and click **OK**. After about 5 minutes, user A can access the table created by user B.
9. Perform the following steps to add the Ranger access permission policy of Hive:
   a. Log in to FusionInsight Manager as a Hive administrator and choose **Cluster** > **Services**. On the page that is displayed, choose **Ranger**. On the displayed page, click the URL next to **Ranger WebUI** to go to the Ranger management page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

b. On the home page, click the component plug-in name in the **HADOOP SQL** area, for example, **Hive**.
c. On the **Access** tab page, click **Add New Policy** to add a Hive permission control policy.
d. In the **Create Policy** dialog box that is displayed, set the following parameters:
   - **Policy Name**: Enter a policy name, for example, **table_test_hive**.
   - **database**: Enter or select the database where the table created by user B is stored, for example, **default**.
   - **table**: Enter or select the table created by user B, for example, **test**.
   - **column**: Enter and select a column, for example, *****.
   - In the **Allow Conditions** area, click **Select User**, select user A, click **Add Permissions**, and select **select**.
   - Click **Add**.

10. Perform the following steps to add the Ranger access permission policy of HDFS:
    a. Log in to FusionInsight Manager as user **rangeradmin** and choose **Cluster** > **Services**. On the page that is displayed, choose **Ranger**. On the displayed page, click the URL next to **Ranger WebUI** to go to the Ranger management page.
    b. On the home page, click the component plug-in name in the **HDFS** area, for example, **hacluster**.
    c. Click **Add New Policy** to add an HDFS permission control policy.
    d. In the **Create Policy** dialog box that is displayed, set the following parameters:
       - **Policy Name**: Enter a policy name, for example, **tablehdfs_test**.
       - **Resource Path**: Set this parameter to the HDFS path where the table created by user B is stored, for example, **/user/hive/warehouse/***Database name*/*Table name*.
       - In the **Allow Conditions** area, select user A for **Select User**, click **Add Permissions** in the **Permissions** column, and select **Read** and **Execute**.
       - Click **Add**.
11. View basic information about the policy in the policy list. After the policy takes effect, user A can view the table created by user B.

11.4.7 How Do I Prevent Kerberos Authentication Expiration?

- Java applications:

  Before connecting to HBase, HDFS, or other big data components, call loginUserFromKeytab() to create a UGI. Then, start a scheduled thread to periodically check whether the Kerberos Authentication expires. Log in to the system again before the Kerberos Authentication expires.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
private static void startCheckKeytabTgtAndReloginJob() {
    //The credential is checked every 10 minutes, and updated before the expiration time.
    ThreadPool.updateConfigThread.scheduleWithFixedDelay(() -> {
        try {
            UserGroupInformation.getLoginUser().checkTGTAndReloginFromKeytab();
            logger.warn("get tgt:{}", UserGroupInformation.getLoginUser().getTGT());
            logger.warn("Check Kerberos Tgt And Relogin From Keytab Finish.");
        } catch (IOException e) {
            logger.error("Check Kerberos Tgt And Relogin From Keytab Error", e);
        }
    }, 0, 10, TimeUnit.MINUTES);
    logger.warn("Start Check Keytab TGT And Relogin Job Success.");
}
```

- Tasks executed in shell client:
  a. Run the **kinit** command to authenticate the user.
  b. Create a scheduled task of the operating system or any other scheduled task to run the **kinit** command to authenticate the user periodically.
  c. Submit jobs to execute big data tasks.
- Spark jobs:

  If you submit jobs using **spark-shell**, **spark-submit**, or **spark-sql**, you can specify **Keytab** and **Principal** in the command to perform authentication and periodically update the login credential and authorization tokens to prevent authentication expiration.

  The following is an example:

  spark-shell --principal spark2x/hadoop.<*System domain name*>@<*System domain name*> **--keytab** *${BIGDATA_HOME}/FusionInsight_Spark2x_XXX/install/FusionInsight-Spark2x-2.4.5/keytab/spark2x/SparkResource/spark2x.keytab* **--master yarn**

11.4.8 How Do I Enable or Disable Kerberos Authentication for an Existing MRS Cluster?

Question:

How do I enable or disable Kerberos authentication for an existing MRS cluster?

Answer:

You cannot enable or disable the Kerberos service after the MRS cluster is created. You need to create an MRS cluster, enable or disable the Kerberos service, and migrate data from the original one.

11.4.9 What Are the Ports of the Kerberos Authentication Service?

Question:

What Are the Ports of the Kerberos Authentication Service?

Answer:

The Kerberos authentication service uses ports 21730 (TCP), 21731 (TCP/UDP), and 21732 (TCP/UDP).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.5 Client Usage

11.5.1 How Do I Disable SASL Authentication for ZooKeeper?

Question:

How Do I Disable ZooKeeper SASL Authentication?

Answer:

1. Log in to FusionInsight Manager.
2. Choose Cluster > Service > ZooKeeper > Configuration > All Configurations.
3. In the navigation tree on the left, choose quorumpeer > Customization and add zookeeper.sasl.disable = false.

   Save the configuration and restart the ZooKeeper service.

11.5.2 What Can I Do If the Error Message "Permission denied" Is Displayed When kinit Is Executed on a Client Outside the MRS Cluster?

Symptom

After the client is installed on a node outside an MRS cluster and the **kinit** command is executed, the following error information is displayed:

```
-bash kinit Permission denied
```

The following error information is displayed when the **java** command is executed:

```
-bash: /xxx/java: Permission denied
```

After running the **ll /***Java installation path***/JDK/jdk/bin/java** command, it is found that the file execution permission is correct.

Fault Locating

Run the **mount | column -t** command to check the status of the mounted partition. It is found that the partition status of the mount point where the Java execution file is located is **noexec**. In the current environment, the data disk where the MRS client is installed is set to **noexec**, that is, binary file execution is prohibited. As a result, Java commands cannot be executed.

Solution

1. Log in to the node where the MRS client is located as user **root**.
2. Remove the configuration item **noexec** of the data disk where the MRS client is located from the **/etc/fstab** file.
3. Run the **umount** command to detach the data disk, and then run the **mount -a** command to remount the data disk.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.5.3 What Should I Do If an Alarm Is Reported Indicating that the Memory Is Insufficient When I Execute a SQL Statement on the ClickHouse Client?

Symptom

The ClickHouse client restricts the memory used by GROUP BY statements. When a SQL statement is executed on the ClickHouse client, the following error information is displayed:

```
Progress: 1.83 billion rows, 85.31 GB (68.80 million rows/s., 3.21 GB/s.)  6%Received exception from server:
Code: 241. DB::Exception: Received from localhost:9000, 127.0.0.1.
DB::Exception: Memory limit (for query) exceeded: would use 9.31 GiB (attempt to allocate chunk of 1048576 bytes), maximum: 9.31 GiB:
(while reading column hits):
```

Solution

- Run the following command before executing an SQL statement on condition that the cluster has sufficient memory:
  ```
  SET max_memory_usage = 128000000000; #128G
  ```
- If no sufficient memory is available, ClickHouse enables you to overflow data to disk to free up the memory: You are advised to set the value of **max_memory_usage** to twice the size of **max_bytes_before_external_group_by**.
  ```
  set max_bytes_before_external_group_by=20000000000; #20G
  set max_memory_usage=40000000000; #40G
  ```
- If the customer has a large amount of data and the entire table is queried, you are advised to query the data by partition or upgrade the specifications of the core node in the cluster.

11.5.4 How Do I Connect to Spark Shell from MRS?

1. Log in to the Master node in the cluster as user **root**.
2. Run the following command to configure environment variables:

   source *Client installation directory*/**bigdata_env**
3. If Kerberos authentication is enabled for the cluster, authenticate the user. If Kerberos authentication is disabled, skip this step.

   Run the **kinit** *MRS cluster user* command.

   Examples:

   - For a machine-machine user, run the **kinit -kt** *user.keytab sparkuser* command.
   - For a human-machine user, run the **kinit** *sparkuser* command.
4. Run the following command to connect to Spark shell:

   spark-shell

11.5.5 How Do I Connect to Spark Beeline from MRS?

1. Log in to the master node in the cluster as user **root**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2. Run the following command to configure environment variables:

   source *Client installation directory*/**bigdata_env**

3. If Kerberos authentication is enabled for the cluster, authenticate the user. If Kerberos authentication is disabled, skip this step.

   Run the **kinit** *MRS cluster user* command.

   Examples:

   - For a machine-machine user, run the **kinit -kt** *user.keytab sparkuser* command.
   - For a human-machine user, run the **kinit** *sparkuser* command.

4. Run the following command to connect to Spark Beeline:

   spark-beeline

5. Run commands on Spark Beeline. For example, create the table **test** in the **obs://mrs-word001/table/** directory.

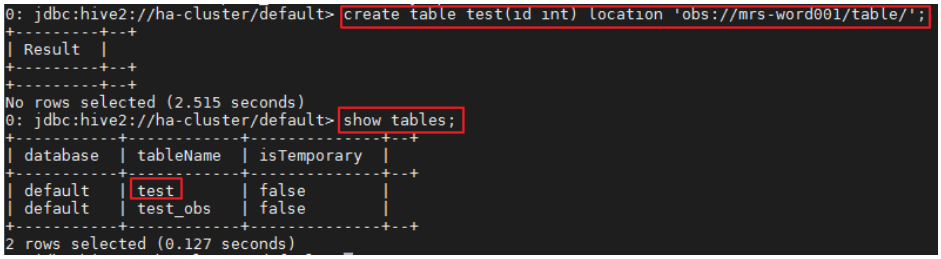
   create table test(id int) location 'obs://mrs-word001/table/';

6. Query all tables.

   show tables;

   If the table **test** is displayed in the command output, OBS is successfully accessed.

   Figure 11-4 Returned table name

7. Press **Ctrl+C** to exit the Spark Beeline.

11.5.6 What Should I Do If the Connection to the ClickHouse Server Fails and Error Code 516 Is Reported?

Symptom

When the **clickhouse client** command is executed to connect to the ClickHouse server, the following error is reported:

```
ClickHouse exception, code: 516, host: 192.168.0.198, port: 8443; Code: 516, e.displayText() = DB::Exception: clickDevelopuser: Authentication failed: password is incorrect or there is no user with such name
```

Fault Locating

The username or password used for connecting the ClickHouse server is incorrect.

Procedure

Enter the correct username or password .

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.6 Component Configurations

11.6.1 Does MRS Support Running Hive on Kudu?

Question:

Does MRS Support Running Hive on Kudu?

Answer:

MRS does not support Hive on Kudu.

Currently, MRS supports only the following two methods to access Kudu:

- Access Kudu through Impala tables.
- Access and operate Kudu tables using the client application.

11.6.2 Does an MRS Cluster Support Hive on Spark?

Question:

Does an MRS Cluster Support Hive on Spark?

Answer:

- Clusters of MRS 1.9.*x* support Hive on Spark.
- Clusters of MRS 3.*x* or later support Hive on Spark.
- You can use Hive on Tez for the clusters of other versions.

11.6.3 Can I Change the IP address of DBService?

Question:

Can I Change the IP address of DBService?

Answer:

The IP address of DBService cannot be changed in the MRS cluster.

11.6.4 What Access Protocols Does Kafka Support?

Question:

What Access Protocols Does Kafka Support?

Answer:

Kafka supports the following access protocols: PLAINTEXT, SSL, SASL_PLAINTEXT, and SASL_SSL.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.6.5 What Python Versions Are Supported by Spark Tasks in an MRS Cluster?

Question:

Which Python Versions Are Supported by Spark Tasks in MRS 3.1.0 Clusters?

Answer:

For MRS 3.1.0 clusters, Python 2.7 or 3.x is recommended for Spark tasks.

11.6.6 What Are the Restrictions on the Storm Log Size in an MRS 2.1.0 Cluster?

Question:

What Are the Restrictions on the Storm Log Size in an MRS 2.1.0 Cluster?

Answer:

In MRS 2.1.0, the size of Storm logs cannot exceed 20 GB. If the size of Storm logs exceeds 20 GB, the logs will be deleted cyclically.

Logs are stored on the system disk, and the space is limited. If you want to store logs for a long time, mount them to a data disk.

11.6.7 How Do I Modify the HDFS fs.defaultFS of an Existing Cluster?

Question:

How do I modify the HDFS NameSpace (fs.defaultFS) of an existing cluster?

Answer:

You can modify or add the HDFS NameSpace (fs.defaultFS) of the cluster by modifying the parameters in the **core-site.xml** and **hdfs-site.xml** files on the client. However, you are not advised to perform this operation on the server.

11.6.8 Can MRS Run Multiple Flume Tasks at a Time?

The Flume client supports multiple independent data flows. You can configure and link multiple sources, channels, and sinks in the **properties.properties** configuration file.

These components can be linked to form multiple data flows.

The following is an example of configuring two data flows in a configuration file:

```
server.sources = source1 source2
server.sinks = sink1 sink2
server.channels = channel1 channel2

#dataflow1
server.sources.source1.channels = channel1
server.sinks.sink1.channel = channel1
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
#dataflow2
server.sources.source2.channels = channel2
server.sinks.sink2.channel = channel2
```

11.6.9 How Do I Change FlumeClient Logs to Standard Logs?

1. Log in to the node where the Flume client is installed.
2. Go to the Flume client installation directory, for example, **/opt/FlumeClient**. Run the following command:

   cd /opt/FlumeClient/fusioninsight-flume-1.9.0/bin
3. Run the **./flume-manage.sh stop force** command to stop FlumeClient.
4. Run the **vi ../conf/log4j.properties** command to open the **log4j.properties** file and change the **flume.root.logger** value to **${flume.log.level},console**.
5. Run the **./flume-manage.sh start force** command to restart FlumeClient.
6. After the modification, check whether the Docker configuration is correct.

11.6.10 Where Are the JAR Files and Environment Variables of Hadoop Stored?

- **hadoopstreaming.jar**: **/opt/share/hadoop-streaming-*** (***** indicates the Hadoop version.)
- The JDK environment variables are stored in **/opt/client/JDK/component_env**.
- The Hadoop environment variables are stored in **/opt/client/HDFS/component_env**.
- The Hadoop client path is **/opt/client/HDFS/hadoop**.

11.6.11 How Do I View HBase Logs?

1. Log in to the Master node in the cluster as user **root**.
2. Run the **su - omm** command to switch to user **omm**.
3. Run the **cd /var/log/Bigdata/hbase/** command to go to the **/var/log/Bigdata/hbase/** directory and view HBase logs.

11.6.12 How Do I Set the TTL for an HBase Table?

- Set the time to live (TTL) when creating a table:

  Create the **t_task_log** table, set the column family to **f**, and set the TTL to **86400** seconds.

  ```
  create 't_task_log',{NAME => 'f', TTL=>'86400'}
  ```
- Set the TTL for an existing table:

  ```
  alter "t_task_log",NAME=>'data',TTL=>'86400' # Set the TTL value for the column family data.
  ```

11.6.13 How Do I Change the Number of HDFS Replicas?

1. Go to the HDFS service configuration page.
   - For MRS 1.8.10 or earlier clusters

     Log in to MRS Manager (see), choose **Services** > **HDFS** > **Service Configuration**, and select **All** from the **Basic** drop-down list.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - For MRS 1.8.10 or later and MRS 2.*x*, click the cluster name. On the cluster details page that is displayed, choose **Components** > **HDFS** > **Service Configuration**, and select **All** from the **Basic** drop-down list.

    NOTE

    If the **Components** tab is unavailable, complete IAM user synchronization first. (On the **Dashboard** page, click **Synchronize** on the right side of **IAM User Sync** to synchronize IAM users.)

  - For MRS cluster version 3.*x* or later:

    Log in to FusionInsight Manager, and choose **Cluster** > *Name of the desired cluster* > **Services** > **HDFS** > **Configurations** > **All Configurations**.

2. Search for **dfs.replication**, change the value (value range: 1 to 16), and restart the HDFS instance.

11.6.14 How Do I Modify the HDFS Active/Standby Switchover Class?

If the **org.apache.hadoop.hdfs.server.namenode.ha.AdaptiveFailoverProxyProvider** class is unavailable when an MRS 3.*x* cluster connects to NameNodes using HDFS, the cause is that the HDFS active/standby switchover class of the cluster is configured improperly. To solve the problem, perform the following operations:

- Method 1: Add the **hadoop-plugins-*xxx*.jar** package to the **classpath** or **lib** directory of your program.

  The **hadoop-plugins-*xxx*.jar** package is stored in the HDFS client directory, for example, **$HADOOP_HOME/share/hadoop/common/lib/hadoop-plugins-*xxx*.jar**.

- Method 2: Change the following HDFS configuration item to the corresponding open source class:

  dfs.client.failover.proxy.provider.hacluster=org.apache.hadoop.hdfs.server.namenode.ha.ConfiguredFailoverProxyProvider

11.6.15 What Data Type in Hive Tables Is Recommended for the Number Type of DynamoDB?

Question:

What is the best type for the number in DynamoDB in Hive tables?

Answer:

smallint is recommended.

11.6.16 Can I Export the Query Result of Hive Data?

Question:

Can I export the query result of Hive data in an MRS cluster?

Answer:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Run the following statement to export the query result of Hive data:

```
insert overwrite local directory "/tmp/out/" row format delimited fields terminated by "\t" select * from table;
```

11.6.17 What Should I Do If an Error Occurs When Hive Runs the beeline -e Command to Execute Multiple Statements?

When Hive uses Beeline to run the **beeline -e "use default;show tables;"** command on an MRS 3.x cluster, the following error is reported:

```
Error while compiling statement: FAILED: ParseException line 1:11 missing EOF at ';' near 'default' (state=42000,code=40000)
```

Solutions:

- Method 1: Replace the **beeline -e " use default;show tables;"** command with **beeline --entirelineascommand=false -e "use default;show tables;"**.
- Method 2:
  a. In the **/opt/Bigdata/client/Hive** directory on the Hive client, change **export CLIENT_HIVE_ENTIRELINEASCOMMAND=true** in the **component_env** file to **export CLIENT_HIVE_ENTIRELINEASCOMMAND=false**.

     Figure 11-5 Changing the **component_env** file

  b. Run the following command to verify the configuration:

     source /opt/Bigdata/client/bigdata_env

     beeline -e "use default;show tables;"

11.6.18 What Do I Do If "over max user connections" Is Displayed When Hue Connects to HiveServer?

Applicable versions: MRS 3.1.0 and earlier

1. Modify the following configuration files on all Hue nodes:

   /opt/Bigdata/FusionInsight_Porter_8.*/install/FusionInsight-Hue-*/hue/apps/beeswax/src/beeswax/models.py

2. Change the configurations in lines 396 and 404.

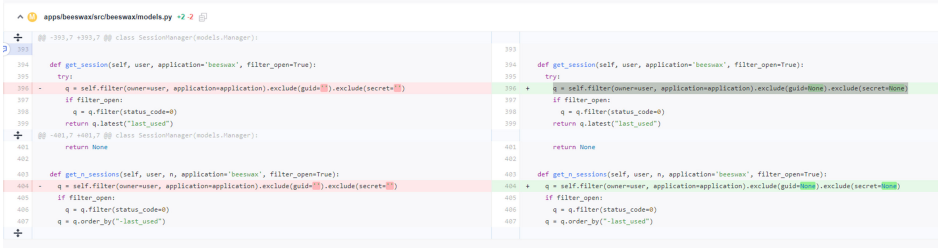
   ```
   q = self.filter(owner=user, application=application).exclude(guid='').exclude(secret='')
   ```

   Modify the content as follows:

   ```
   q = self.filter(owner=user, application=application).exclude(guid=None).exclude(secret=None)
   ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 11-6 Modifying the Hue configuration file



11.6.19 How Do I View MRS Hive Metadata?

- Hive metadata is stored in the GaussDB database of the MRS cluster. You can perform the following steps to view the metadata:
  a. Log in to the DBServer active node as user **root**.
  b. Log in to the Hive database:
     su - omm
     source $DBSERVER_HOME/.dbservice_profile
     gsql -p 20051 -U *USER* **-W** *PASSWD* **-d hivemeta**
  c. View all metadata tables in the Hive metadata database:
     \d+
  d. Run **q** to exit the page for viewing the metadata table.
  e. View data in the metadata table:
     select * from *Table name*;

  NOTE

  To query the IP address of the active DBServer node, log in to FusionInsight Manager, choose **Cluster** > **Services** > **DBService**, and click the **Instances** tab.
- If Hive metadata is stored in an external relational database, perform the following steps:
  a. On the cluster **Dashboard** page, click **Manage** on the right of **Data Connection**.
  b. On the displayed page, obtain the value of **Data Connection ID**.
  c. On the MRS console, click **Data Connections**.
  d. In the data connection list, locate the data connection based on the data connection ID.
  e. Click **Edit** in the **Operation** column of the data connection.
     The **RDS Instance** and **Database** indicate the relational database in which the Hive metadata is stored.

11.6.20 How Do I Reset Kafka Data?

You can reset Kafka data by deleting Kafka topics.

- Delete a topic: **kafka-topics.sh --delete --zookeeper** *ZooKeeper Cluster service IP address*:**2181/kafka --topic** ***topicname***
- Query all topics: **kafka-topics.sh --zookeeper** *ZooKeeper cluster service IP address*:**2181/kafka --list**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

After the deletion command is executed, empty topics will be deleted immediately. If a topic has data, the topic will be marked for deletion and will be deleted by Kafka later.

11.6.21 What Should I Do If the Error Message "Not Authorized to access group XXX" Is Displayed When Kafka Topics Are Consumed?

This issue is caused by the conflict between the Ranger authentication and ACL authentication of a cluster. If a Kafka cluster uses ACL for permission access control and Ranger authentication is enabled for the Kafka component, all authentications of the component are managed by Ranger. The permissions set by the original authentication plug-in are invalid. As a result, ACL authorization does not take effect. You can disable Ranger authentication of Kafka and restart the Kafka service to rectify the fault. The procedure is as follows:

1. Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Kafka**.
2. In the upper right corner of the **Dashboard** page, click **More** and choose **Disable Ranger**. In the displayed dialog box, enter the password and click **OK**. After the operation is successful, click **Finish**.
3. In the upper right corner of the **Dashboard** page, click **More** and choose **Restart Service** to restart the Kafka service.

11.6.22 What Compression Algorithms Does Kudu Support?

Question:

What compression algorithms does Kudu support?

Answer:

It supports **Snappy**, **LZ4**, and **zlib**. **LZ4** is used by default.

11.6.23 How Do I View Kudu Logs?

1. Log in to the Master node in the cluster.
2. Run the **su - omm** command to switch to user **omm**.
3. Run the **cd /var/log/Bigdata/kudu/** command to go to the **/var/log/Bigdata/kudu/** directory and view Kudu logs.

11.6.24 How Do I Handle the Kudu Service Exceptions Generated During Cluster Creation?

Viewing the Kudu Service Exception Logs

1. Log in to the MRS console.
2. Click the name of the cluster.
3. On the page displayed, choose **Components** > **Kudu** > **Instances** and locate the IP address of the abnormal instance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the **Components** tab is unavailable, complete IAM user synchronization first. (On the **Dashboard** page, click **Synchronize** on the right side of **IAM User Sync** to synchronize IAM users.)

4. Log in to the node where the abnormal instance resides, and view the Kudu log.

```
cd /var/log/Bigdata/Kudu
[root@node-master1AERu kudu]# ls
healthchecklog  runninglog  startlog
```

You can find the Kudu health check logs in the **healthchecklog** directory, the startup logs in the **startlog** directory, and the Kudu process run logs in the **runninglog** directory.

```
[root@node-master1AERu logs]# pwd
/var/log/Bigdata/kudu/runninglog/master/logs
[root@node-master1AERu logs]# ls -al
kudu-master.ERROR   kudu-master.INFO   kudu-master.WARNING
```

Run logs are classified into three types: ERROR, INFO, and WARNING. Each type of run logs is recorded in the corresponding file. You can run the **cat** command to view run logs of each type.

Handling Kudu Service Exceptions

The **/var/log/Bigdata/kudu/runninglog/master/logs/kudu-master.INFO** file contains the following error information:

```
"Unable to init master catalog manager: not found: Unable to initialize catalog manager: Failed to initialize sys tables async: Unable to load consensus metadata for tablet 00000000000000000000: xxx"
```

If this exception occurs when the Kudu service is installed for the first time, the KuduMaster service is not started. Data inconsistency causes the startup failure. To solve the problem, perform the following steps to clear the data directories and restart the Kudu service. If the Kudu service is not installed for the first time, clearing the data directories will cause data loss. In this case, migrate data and clear the data directory.

1. Search for the data directories **fs_data_dir**, **fs_wal_dir**, and **fs_meta_dir**.

   find /opt -name master.gflagfile

   cat /opt/Bigdata/FusionInsight_Kudu_*/*_KuduMaster/etc/master.gflagfile | grep fs_

2. On the cluster details page, choose **Components** > **Kudu** and click **Stop Service**.
3. Clear the Kudu data directories on all KuduMaster and KuduTserver nodes. The following command uses two data disks as an example.

   rm -Rvf /srv/Bigdata/data1/kudu, rm -Rvf /srv/Bigdata/data2/kudu

4. On the cluster details page, choose **Components** > **Kudu** and choose **More** > **Restart Service**.
5. Check the Kudu service status and logs.

11.6.25 How Do I Configure Other Data Sources on Presto?

This topic uses the MySQL data source as an example.

- For MRS 1.*x* and 3.*x* clusters, do the following:

  a. Log in to the MRS management console.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

b. Click the name of the cluster to go to its details page.

c. Click the **Components** tab and then **Presto** in the component list. On the page that is displayed, click the **Configurations** tab then the **All Configurations** sub-tab.

d. On the Presto configuration page that is displayed, find **connector-customize**.

e. Set **Name** and **Value** as follows:

Name: **mysql.connector.name**

Value: **mysql**

f. Click the plus sign (+) to add three more fields and set **Name** and **Value** according to the table below. Then click **Save**.

Table 11-1 Parameter description

Name	Value	Description
mysql.connection-url	jdbc:mysql://xxx.xxx.xxx.xxx:3306	Database connection pool
mysql.connection-user	xxxx	Database username
mysql.connection-password	xxxx	Database password

g. Restart the Presto service.

h. Run the following command to connect to the Presto Server of the cluster:

presto_cli.sh --*krb5-config-path* {krb5.conf path} *--krb5-principal* {User principal} *--krb5-keytab-path* {user.keytab path} *--user* {presto username}

i. Log in to Presto and run the **show catalogs** command to check whether the data source list of Presto can be queried.

```
[root@node-master2uoHG bin]# ./presto_cli.sh
--server http://15[redacted]20
        show catalogs;
 Catalog
---------
 hive
 jmx
 mysql
 system
 tpcds
 tpch
(6 rows)

Query 20220422_121338_00002_ra2vb, FINISHED, 3 nodes
Splits: 53 total, 53 done (100.00%)
0:00 [0 rows, 0B] [0 rows/s, 0B/s]
```

Run the **show schemas from mysql** command to query the MySQL database.

- For MRS 2.*x* clusters, do the following:

a. Create the **mysql.properties** configuration file containing the following content:

```
connector.name=mysql
connection-url=jdbc:mysql://IP address of MySQL:3306
connection-user=Username
connection-password=Password
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- The IP Address of MySQL is that of the MySQL instance, which must be able to communicate with the MRS cluster.
- The username and password are those used to log in to the MySQL database.
- Configuration files containing authentication passwords pose security risks. Delete such files after configuration or store them securely.

b. Upload the configuration file to the **/opt/Bigdata/MRS_Current/1_14_Coordinator/etc/catalog/** directory on the Master node (where the Coordinator instance resides) and the **/opt/Bigdata/MRS_Current/1_14_Worker/etc/catalog/** directory on the Core node (depending on the actual directory in the cluster), and change the file owner group to **omm:wheel**.

c. Restart the Presto service.

11.6.26 How Do I Specify a Log Path When Submitting a Task in an MRS Storm Cluster?

Question:

How do I specify a log path when submitting a task in an MRS storm cluster?

Answer:

You can open the **/opt/Bigdata/MRS_*XXX*/1_*XX*_Supervisor/etc/worker.xml** file on the streaming Core node of MRS, set the value of **filename** in the file to the path, and restart the corresponding instance on Manager.

NOTE

You are advised not to modify the default log configuration of MRS. Otherwise, the log system may be abnormal.

11.6.27 How Do I Check the ResourceManager Configuration of Yarn?

This section applies only to versions earlier than MRS 3.*x*.

Step 1 Log in to MRS Manager and choose **Services** > **Yarn** > **Instance**.

Step 2 Click the names of the two ResourceManager nodes, choose **More** > **Synchronize Configuration**, and deselect **Restart services or instances whose configurations have expired**. .

Step 3 Click **Yes** to synchronize the configuration.

Step 4 Log in to the Master nodes as user **root**.

Step 5 Run the **cd /opt/Bigdata/MRS_Current/*_*_ResourceManager/etc_UPDATED/** command to go to the **etc_UPDATED** directory.

Step 6 Run the **grep '\.queues' capacity-scheduler.xml -A2** command to display all configured queues and check whether the queues are consistent with those displayed on Manager.

root-default is hidden on the Manager page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
[omm@node-master1IlZA etc]$
[omm@node-master1IlZA etc]$ grep '\.queues' capacity-scheduler.xml  -A2
<name>yarn.scheduler.capacity.root.queues</name>
<value>default,root-default,launcher-job,test1,test2,test3,test4</value>
</property>
[omm@node-master1IlZA etc]$
[omm@node-master1IlZA etc]$
```

Step 7 Run the **grep '\.capacity</name>' capacity-scheduler.xml -A2** command to display the value of each queue and check whether the value of each queue is the same as that displayed on Manager. Check whether the sum of the values configured for all queues is **100**.

- If the sum is **100**, the configuration is correct.
- If the sum is not **100**, the configuration is incorrect. Perform the following steps to rectify the fault.

Step 8 Log in to MRS Manager, and select **Hosts**.

Step 9 Determine the active Master node. The host name of the active Master node starts with a solid pentagon.

Step 10 Log in to the active Master node as user **root**.

Step 11 Run the **su - omm** command to switch to user **omm**.

Step 12 Run the **sh /opt/Bigdata/om-0.0.1/sbin/restart-controller.sh** command to restart the controller when no operation is being performed on Manager.

Restarting the controller will not affect the big data component services.

Step 13 Repeat **Step 1** to **Step 7** to synchronize ResourceManager configurations and check whether the configurations are correct.

If the latest configuration has not been loaded after the configuration synchronization is complete, a message will be displayed on the Manager page indicating that the configuration has expired. However, this will not affect services. The latest configuration will be automatically loaded when the component restarts.

----End

11.6.28 How Do I Modify the allow_drop_detached Parameter of ClickHouse?

Step 1 Log in to the node where the ClickHouse client is located as user **root**.

Step 2 Run the following commands to go to the client installation directory and set the environment variables:

cd /opt/*Client installation directory*

source bigdata_env

Step 3 If Kerberos authentication is enabled for the cluster, run the following command to authenticate the user. If Kerberos authentication is disabled, skip this step.

kinit *MRS cluster user*

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

> **NOTE**
>
> The user must have the ClickHouse administrator permissions.

Step 4 Run the **clickhouse client --host** *192.168.42.90* **--secure -m** command, in which *192.168.42.90* indicates the IP address of the ClickHouseServer instance node. The command output is as follows:

```
[root@server-2110082001-0017 hadoopclient]# clickhouse client --host 192.168.42.90 --secure -m
ClickHouse client version 21.3.4.25.
Connecting to 192.168.42.90:21427.
Connected to ClickHouse server version 21.3.4 revision 54447.
```

Step 5 Run the following command to set the value of the **allow_drop_detached** parameter, for example, **1**:

set allow_drop_detached=1;

Step 6 Run the following command to query the value of the **allow_drop_detached** parameter:

SELECT * FROM system.settings WHERE name = 'allow_drop_detached';



Step 7 Run the **q;** command to exit the ClickHouse client.

----End

11.6.29 How Do I Add a Policy for Periodically Deleting Large ClickHouse System Table Logs?

Symptom

ClickHouse logs generated in the system table is too large. Deleting logs at a time takes a long time.

Handling Procedure

- For MRS 3.3.0 and later versions, on FusionInsight Manager, choose **Cluster** > **Services** > **ClickHouse** > **Configurations** > **All Configurations**. You can set **query_log.ttl** and **trace_log.ttl** to configure the data clearing period, the default value is 90 days.
- In versions earlier than MRS 3.3.0, run the following statement on the ClickHouse client to modify the TTL of a table in ClickHouse:

  alter table *system.Table name* **modify** *TTL event_date* **+ INTERVAL** *Number of days for storing data* **day;**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

CAUTION

- This statement is used only on a SQL node to configure the TTL of the system table. If the TTL needs to be configured on all nodes, run this statement on each node. Do not run the statement with **on cluster** on each node. Otherwise the ClickHouse will keep running.
- Run the preceding statement during off-peak hours. This operation may take a long time if there is a large amount of data.

11.7 Cluster Management

11.7.1 How Do I View All Clusters?

All MRS clusters are displayed on the **Clusters** page of the MRS management console. You can view clusters in different status.

- **Active Clusters**: all clusters except clusters in **Failed** and **Terminated** states.
- **Cluster History**: clusters in the **Deleted** state. Only the clusters deleted within the last six months are displayed. If you want to view clusters deleted more than six months ago, contact technical support engineers.
- **Failed Tasks**: tasks in **Failed** state. The failed tasks include the following:
  - Tasks failed to create clusters
  - Tasks failed to terminate clusters
  - Tasks failed to scale out clusters
  - Tasks failed to scale in clusters

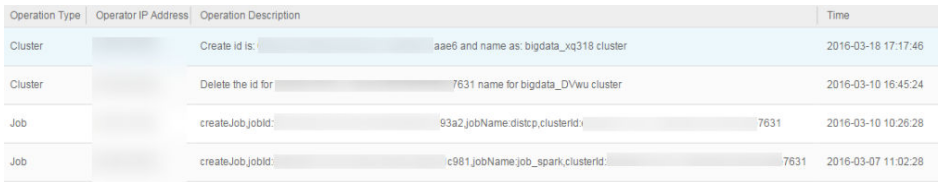
11.7.2 How Do I View MRS Operation Logs?

You can view operation logs of clusters and jobs on the **Operation Logs** page. The MRS operation logs record the following operations:

- Cluster operations
  - Create, terminate, and scale out or in clusters
  - Create directories and delete directories or files
- Job operations: Create, stop, and delete jobs
- Data operations: IAM user tasks, add users, and add user groups

Figure 11-7 shows the operation logs.

Figure 11-7 Log information



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.7.3 How Do I View MRS Cluster Configuration Information?

- After a cluster is created, click the cluster name. On the page displayed, you can view basic configuration information about the cluster.
- The instance specifications and node capacity determine the data analysis and processing capability of the cluster.
- Higher instance specifications and larger capacity enable faster data processing at a higher cost.
- On the **Dashboard** page, click **Access Manager**.
- On displayed MRS Manager, you can view and handle alarms, and modify cluster configurations.

11.7.4 How Do I Add Components to an MRS Cluster?

Question: How do I add services to an existing MRS cluster?

Answer:

You cannot install new components for a created cluster whose version is MRS 3.1.0 or earlier. If you want to use a component that is not installed in your cluster, you need to create a cluster and select the required component.

You can add components to MRS 3.1.2-LTS.3 or later custom clusters.

11.7.5 How Do I Cancel Message Notifications for MRS Cluster Alarms?

1. Log in to the MRS console.
2. Click the name of the cluster.
3. On the page displayed, choose **Alarms** > **Notification Rules**.
4. Locate the row that contains the rule you want to modify, click **Edit** in the **Operation** column, and deselect the alarm or event severity levels.
5. Click **OK**.

11.7.6 Why Is the Resource Pool Memory Displayed in the MRS Cluster Smaller Than the Actual Cluster Memory?

Question:

Why is the resource pool memory displayed in the MRS cluster smaller than the actual cluster memory?

Answer:

In an MRS cluster, MRS allocates 50% of the cluster memory to YARN by default. You manage YARN nodes logically by resource pool. Therefore, the total memory of the resource pool displayed in the cluster is only 50% of the total memory of the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.7.7 What Is the Python Version Installed for an MRS Cluster?

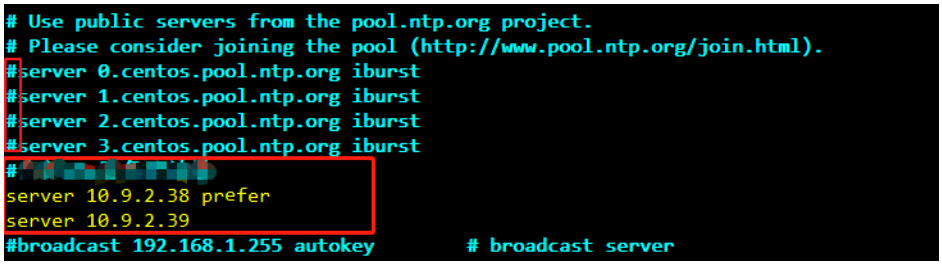
Log in to a Master node as user **root** and run the **Python3** command to query the Python version.

11.7.8 What Can I Do If the Time Information of an MRS Cluster Node Is Incorrect?

- If the time on a node inside the cluster is incorrect, log in to the node and rectify the fault from 2.
- If the time on a node inside the cluster is different from that on a node outside the cluster, log in to the node and rectify the fault from 1.

1. Run the **vi /etc/ntp.conf** command to edit the NTP client configuration file, add the IP addresses of the master node in the MRS cluster, and comment out the IP address of other servers.
   server *master1_ip* prefer
   server *master2_ip*

   Figure 11-8 Adding the master node IP addresses

2. Run the **service ntpd stop** command to stop the NTP service.
3. Run the **/usr/sbin/ntpdate** *IP address of the active master node* command to manually synchronize time.
4. Run the **service ntpd start** or **systemctl restart ntpd** command to start the NTP service.
5. Run the **ntpstat** command to check the time synchronization result:

11.7.9 What Are the Differences and Relationships Between the MRS Management Console and MRS Manager?

Question:

What Are the Differences and Relationships Between the MRS Management Console and MRS Manager?

Answer:

You can access Manager from the MRS management console.

Manager is classified as MRS Manager and FusionInsight Manager.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- MRS Manager is the manager page of MRS 2.*x* or earlier clusters.
- FusionInsight Manager is the manager page of MRS 3.*x* or later clusters.

The following table lists the differences and relationships between the management console and FusionInsight Manager.

Common Operation	MRS Console	FusionInsight Manager
Changing subnets, adding security group rules, controlling OBS permissions, managing agencies, and synchronizing IAM users	Supported	Not supported
Adding node groups, scaling out, scaling in, and upgrading specifications	Supported	Not supported
Isolating hosts, starting all roles, and stopping all roles	Supported	Supported
Downloading the client, starting services, stopping services, and perform rolling restart of services	Supported	Supported
Viewing the instance status of services, configuring parameters, and synchronizing configurations	Supported	Supported
Viewing cleared alarms and events	Supported	Supported
Viewing the alarm help	Supported	Supported
Setting thresholds for threshold alarms	Not supported	Supported
Adding message subscription specifications	Supported	Not supported
Managing files	Supported	Not supported
Managing jobs	Supported	Not supported
Managing tenants	Supported	Supported
Managing tags	Supported	Not supported

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Common Operation	MRS Console	FusionInsight Manager
Setting permissions (adding and deleting users, user groups, and roles)	Not supported	Supported
Backing up component data and restoring data	Not supported	Supported
Logging metadata operation audit	Not supported	Supported
Monitoring resources	Supported	Supported

11.7.10 How Do I Disable the Firewall Service on Nodes in an MRS Cluster?

Question:

How do I disable the firewall service on nodes in an MRS cluster?

Answer:

Step 1 Log in to each node of a cluster as user **root**.

Step 2 Check whether the firewall service is started.

For example, to check the firewall status on EulerOS, run the **systemctl status firewalld.service** command.

Step 3 Stop the firewall service.

For example, to stop the firewall service on EulerOS, run the **systemctl stop firewalld.service** command.

----End

11.7.11 How Do I Switch the Login Mode of a Node in an MRS Cluster?

Question:

How do I switch the login mode (password or key pair) of a node in an MRS cluster?

Answer:

Login mode cannot be changed. You can select the login mode when creating the cluster, but cannot change the login mode after you create the cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.7.12 How Do I Access an MRS Cluster from a Node Outside the Cluster?

Creating a Linux ECS Outside the Cluster to Access the MRS Cluster

Step 1 Create an ECS outside the cluster.

Set **AZ**, **VPC**, and **Security Group** of the ECS to the same values as those of the cluster to be accessed.

Step 2 On the VPC management console, apply for an EIP and bind it to the ECS.

Step 3 Configure security group rules for the cluster.

1. On the **Dashboard** tab page, click **Add Security Group Rule**. In the **Add Security Group Rule** dialog box that is displayed, click **Manage Security Group Rule**.
2. Click the **Inbound Rules** tab, and click **Add Rule**. In the **Add Inbound Rule** dialog box, configure the IP address of the ECS and enable all ports.
3. After the security group rule is added, you can download and install the client on the ECS..
4. Use the client.

   Log in to the client node as the client installation user and run the following command to switch to the client directory:

   cd /opt/hadoopclient

   Run the following command to load environment variables:

   source bigdata_env

   If Kerberos authentication is enabled for the cluster, run the following command to authenticate the user. If Kerberos authentication is disabled for the current cluster, authentication is not required.

   kinit *MRS cluster user*

   Example:

   kinit admin

   Run the client command of a component.

   Example:

   Run the following command to view files in the HDFS root directory:

   hdfs dfs -ls /

   ```
   Found 15 items
   drwxrwx--x   - hive       hive             0 2021-10-26 16:30 /apps
   drwxr-xr-x   - hdfs       hadoop           0 2021-10-18 20:54 /datasets
   drwxr-xr-x   - hdfs       hadoop           0 2021-10-18 20:54 /datastore
   drwxrwx---+  - flink      hadoop           0 2021-10-18 21:10 /flink
   drwxr-x---   - flume      hadoop           0 2021-10-18 20:54 /flume
   drwxrwx--x   - hbase      hadoop           0 2021-10-30 07:31 /hbase
   ...
   ```

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.7.13 Does Kafka Topic Monitoring in MRS Clusters Support Alarm Reporting?

Question:

Can Kafka topic monitoring in MRS streaming clusters send alarms?

Answer:

The Kafka topic monitoring cannot send notifications by email or SMS message. Currently, you can view alarm information on Manager.

11.7.14 How Do I View Running Resource Queues When an Alarm Indicating Insufficient YARN Queue Resources Is Generated?

Question:

Where can I view the running resource queues when ALM-18022 Insufficient Yarn Queue Resources is generated?

Answer:

Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Yarn**. In the navigation pane on the left, choose **ResourceManager(Active)** and log in to the native Yarn page.

For details, see the online help.

11.7.15 How Do I Understand the Multi-Level Chart Statistics in the HBase Operation Requests Metric?

Question:

How Do I Understand the Multi-Level Chart Statistics in the HBase Operation Requests Metric?

Answer:

The following uses the **Operation Requests on RegionServers** monitoring item as an example:

1. Log in to FusionInsight Manager and choose **Cluster** > **Services** > **HBase** > **Resource**. On the displayed page, you can view the **Operation Requests on RegionServers** chart. If you click **all**, the top 10 RegionServers ranked by the total number of operation requests in the current cluster are displayed, the statistics interval is 5 minutes.
2. Click a point in the chart. A level-2 chart is displayed, showing the number of operation requests of all RegionServers in the past 5 minutes.
3. Click an operation statistics bar chart. A level-3 chart is displayed, showing the distribution of operations in each region within the period.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4. Click a region name. The distribution chart of operations performed every 5 minutes in the last 12 hours is displayed. You can view the number of operations performed in the period.

11.8 Node Management

11.8.1 Can I Change MRS Cluster Nodes on the MRS Console?

You cannot change MRS cluster nodes on the MRS console. You are also advised not to change MRS cluster nodes on the ECS console.

Manually stopping or deleting an ECS, modifying or reinstalling the ECS OS, or modifying ECS specifications for a cluster node on the ECS console will affect the cluster stability.

If an ECS is deleted, the ECS OS is modified or reinstalled, or the ECS specifications are modified on the ECS console, MRS will automatically identify and delete the node.

You can log in to the MRS console and restore the deleted node through scale-out. Do not perform operations on the nodes that are being scaled out.

11.8.2 How Do I Query the Startup Time of an MRS Node?

Step 1 Log in to the MRS console, click a cluster name to go to its details page.

Step 2 Click the **Nodes** tab and view the IP address of the node.

Step 3 Log in to the target node and run the following command to query the startup time:

date -d "$(awk -F. '{print $1}' /proc/uptime) second ago" +"%Y-%m-%d %H:%M:%S"

```
[root@server-2110082001-0018 ~]#date -d "$(awk -F. '{print $1}' /proc/uptime) second ago" +"%Y-%m-%d %H:%M:%S"
2021-12-13 15:56:23
```

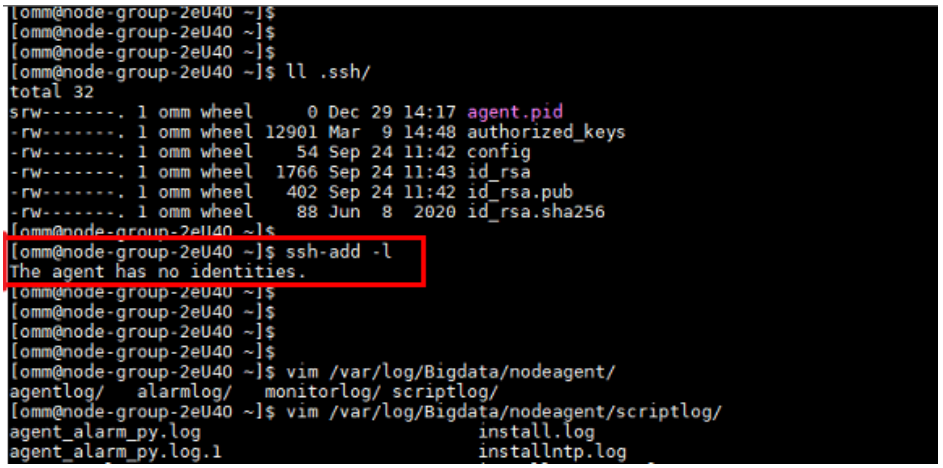
----End

11.8.3 What Do I Do If Trust Relationships Between Nodes Are Abnormal?

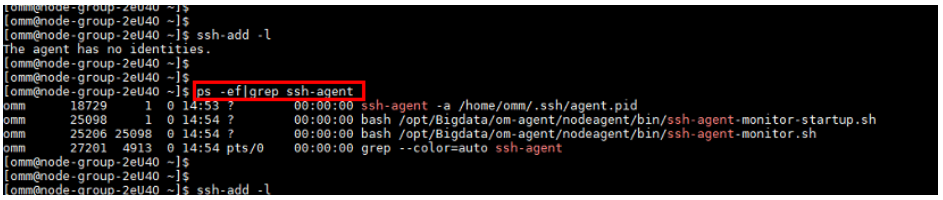
If "ALM-12066 Inter-Node Mutual Trust Fails" is reported on Manager or there is no SSH trust relationship between nodes, rectify the fault by performing the following operations:

1. Run the **ssh-add -l** command on both nodes of the trusted cluster to check whether there are identities.

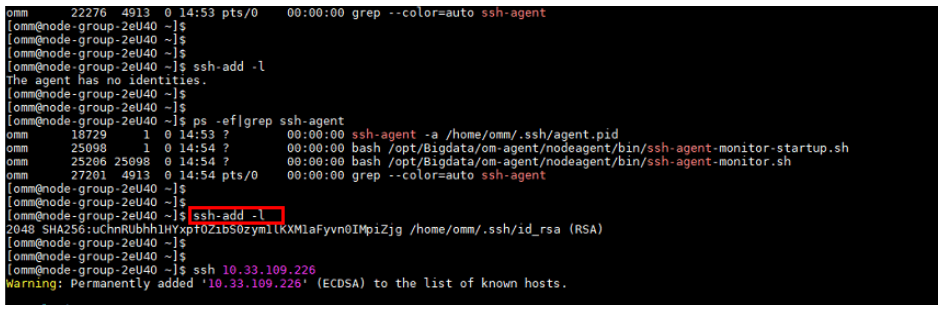
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.



2. If no identities are displayed, run the **ps -ef|grep ssh-agent** command to find the ssh-agent process, kill the process, and wait for the process to automatically restart.



3. Run the **ssh-add -l** command to check whether the identities have been added. If yes, manually run the **ssh** command to check whether the trust relationship is normal.



4. If identities exist, check whether the **authorized_keys** file in the **/home/omm/.ssh** directory contains the information in the **id_rsa.pub** file in the **/home/omm/.ssh** of the peer node. If no, manually add the information about the peer node.
5. Check whether the permissions on the files in **/home/omm/.ssh** directory are correct.
6. Check the **/var/log/Bigdata/nodeagent/scriptlog/ssh-agent-monitor.log** file.
7. If the **omm** user directory has been deleted, contact MRS support personnel.

11.8.4 Can Sudo Logs of Nodes in an MRS Cluster Be Cleared?

Question:

Can Sudo Logs of Nodes in an MRS Cluster Be Cleared?

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Answer:

The sudo log file on the nodes in the MRS cluster records the operations performed by user **omm** to facilitate fault locating. You can clear the file as you need.

You can delete the logs in earlier days to release storage space.

1. If the log file is large, add the log file directory to **/etc/logrotate.d/syslog** to enable periodical deletion.

   Run **sed -i '3 a/var/log/sudo/sudo.log' /etc/logrotate.d/syslog**.

2. Set the maximum number and size of logs in **/etc/logrotate.d/syslog**. If the threshold is exceeded, the logs will be automatically deleted. By default, logs are aged based on the size and number of archived logs. You can use **size** and **rotate** to limit the size and number of archived logs, respectively. If required, you can also add **daily**/**weekly**/**monthly** to specify how often the logs are cleared.

11.8.5 Does an MRS Cluster Support System Reinstallation?

Question:

Does an MRS cluster support system reinstallation?

Answer:

The nodes in the MRS cluster use initial ECS images as their OSs, which cannot be reinstalled.

11.8.6 Can I Change the OS of an MRS Cluster?

Question:

Can I change the OS of an MRS cluster?

Answer:

The node OS in the MRS cluster is based on the initial ECS image and cannot be changed.

11.9 Component Management

11.9.1 Can I Delete Components Installed in an MRS Cluster?

Question:

Can I delete components from an existing MRS cluster?

Answer:

You cannot delete any component from a cluster in MRS 3.1.0 and earlier versions. If a component is not required, log in to MRS Manager and stop the component on the **Services** page.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.9.2 How Do I View the Configuration File Directory of Each Component?

The configuration file paths of commonly used components are as follows:

Component	Configuration File Directory
ClickHouse	*Client installation directory*/**ClickHouse/clickhouse/config**
Flink	*Client installation directory*/**Flink/flink/conf**
Flume	*Client installation directory/***fusioninsight-flume-***xxx*/**conf**
HBase	*Client installation directory*/**HBase/hbase/conf**
HDFS	*Client installation directory*/**HDFS/hadoop/etc/hadoop**
Hive	*Client installation directory*/**Hive/config**
Hudi	*Client installation directory*/**Hudi/hudi/conf**
Kafka	*Client installation directory*/**Kafka/kafka/config**
Loader	• *Client installation directory*/**Loader/loader-tools-xxx/loader-tool/conf** • *Client installation directory*/**Loader/loader-tools-xxx/schedule-tool/conf** • *Client installation directory*/**Loader/loader-tools-xxx/shell-client/conf** • *Client installation directory*/**Loader/loader-tools-xxx/sqoop-shell/conf**
Oozie	*Client installation directory*/**Oozie/oozie-client-***xxx*/**conf**
Spark2x	*Client installation directory/Spark2x/spark/conf*
Yarn	*Client installation directory/Yarn/config*
ZooKeeper	*Client installation directory/Zookeeper/zookeeper/conf*

11.9.3 Will Upper-Layer Services Be Affected If the Hive Service Status Is Partially Healthy?

Question

Will Upper-Layer Services Be Affected If the Hive Service Status Is Partially Healthy?

Answer:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

If the cluster service status is **Subhealthy**, some enhancements of the service cannot work properly.

Log in to FusionInsight Manager to check the status of other services in the cluster. In this example, the HBase service is stopped. If HBase is installed but is abnormal, Hive, Spark, and Loader are in the **Subhealthy** state.

In this case, upper-layer services depending on the Hive service are not affected. After you restart HBase, the Hive service recovers.

11.9.4 How Do I Obtain the IP Address and Port Number of a ZooKeeper Instance in the MRS Cluster?

You can obtain the IP address and port number of a ZooKeeper instance through the MRS console or FusionInsight Manager.

Method 1: Obtaining the IP address and port number of a ZooKeeper through the MRS console

1. On the **Dashboard** page, click **Synchronize** on the right side of **IAM User Sync** to synchronize IAM users.
2. Click the **Components** tab and choose **ZooKeeper**. On the displayed page, click **Instances** to view the business IP address of a ZooKeeper instance.
3. Click the **Service Configuration** tab. On the displayed page, search for the **clientPort** parameter to view the port number of the ZooKeeper instance.

Method 2: Obtaining the IP address and port number of a ZooKeeper through FusionInsight Manager

1. Log in to FusionInsight Manager.
2. Perform the following operations to obtain the IP address and port number of a ZooKeeper instance.
   - For clusters of MRS 3.*x* or earlier
     i. Choose **Services** > **ZooKeeper**. On the displayed page, click the **Instance** tab to view the business IP address of a ZooKeeper instance.
     ii. Click the **Service Configuration** tab. On the displayed page, search for the **clientPort** parameter to view the port number of the ZooKeeper instance.
   - For clusters of MRS 3.*x* or later
     i. Choose **Cluster** > **Services** > **ZooKeeper**. On the displayed page, click the **Instance** tab to view the business IP address of a ZooKeeper instance.
     ii. Click the **Configurations** tab. On the displayed page, search for the **clientPort** parameter to view the port number of the ZooKeeper instance.

11.10 Job Management

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.1 What Types of Spark Jobs Can Be Submitted in a Cluster?

Question:

What Types of Spark Jobs Can Be Submitted in a Cluster?

Answer:

MRS clusters support Spark jobs submitted in Spark, Spark Script, or Spark SQL mode.

11.10.2 How Do I Enable Different Service Programs to Use Different Yarn Queues?

Question:

How Do I Enable Different Service Programs to Use Different Yarn Queues?

Answer:

Create a tenant on FusionInsight Manager and bind different cluster service users to different tenants.

Procedure

Step 1 Log in to FusionInsight Manager and choose **Tenant Resources**.

Step 2 In the tenant list on the left, select a parent tenant and click . On the page for adding a sub-tenant, set properties for the sub-tenant according to **Table 11-2**.

Table 11-2 Sub-tenant parameters

Parameter	Description
Cluster	Cluster to which the parent tenant belongs.
Parent Tenant	Name of the parent tenant.
Name	• Name of the current tenant. The value consists of 3 to 50 characters, including digits, letters, and underscores (_). • Plan a sub-tenant name based on service requirements. The name cannot be the same as that of a role, HDFS directory, or Yarn queue that exists in the current cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Tenant Type	Whether the tenant is a leaf tenant. • When **Leaf Tenant** is selected, the current tenant is a leaf tenant and no sub-tenant can be added. • When **Non-leaf Tenant** is selected, the current tenant is not a leaf tenant and sub-tenants can be added to the current tenant. However, the tenant depth cannot exceed 5 levels.
Computing Resource	Dynamic compute resources for the current tenant. • When **Yarn** is selected, the system automatically creates a queue in YARN and the queue is named the same as the sub-tenant name. – A leaf tenant can directly submit jobs to the queue. – A non-leaf tenant cannot directly submit jobs to the queue. However, YARN adds an extra queue (hidden) named **default** for the non-leaf tenant to record the remaining resource capacity of the tenant. Actual jobs do not run in this queue. • If **Yarn** is not selected, the system does not automatically create a queue.
Default Resource Pool Capacity (%)	Percentage of computing resources used by the current tenant. The base value is the total resources of the parent tenant.
Default Resource Pool Max Capacity (%)	Maximum percentage of computing resources used by the current tenant. The base value is the total resources of the parent tenant.
Storage Resource	Storage resources for the current tenant. • When **HDFS** is selected, the system automatically creates a folder named after the sub-tenant in the HDFS parent tenant directory. • When **HDFS** is not selected, the system does not automatically allocate storage resources.
Quota	Quota for files and directories.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Parameter	Description
Space Quota	Quota for the HDFS storage space used by the current tenant. • If the unit is set to **MB**, the value ranges from **1** to **8796093022208**. If the unit is set to **GB**, the value ranges from **1** to **8589934592**. • This parameter indicates the maximum HDFS storage space that can be used by the tenant, but not the actual space used. • If its value is greater than the size of the HDFS physical disk, the maximum space available is the full space of the HDFS physical disk. • If this quota is greater than the quota of the parent tenant, the actual storage space does not exceed the quota of the parent tenant.
Storage Path	HDFS storage directory for the tenant. • The system automatically creates a folder named after the sub-tenant name in the directory of the parent tenant by default. For example, if the sub-tenant is **ta1s** and the parent directory is **/tenant/ta1**, the storage path for the sub-tenant is then **/tenant/ta1/ta1s**. • The storage path is customizable in the parent directory.
Description	Description of the current tenant.

NOTE

Roles, computing resources, and storage resources are automatically created when tenants are created.

- The new role has permissions on the computing and storage resources. This role and its permissions are automatically controlled by the system and cannot be manually managed by choosing **System** > **Permission** > **Role**. The role name is in the format of *Tenant name_Cluster ID*. The ID of the first cluster is not displayed by default.
- When using this tenant, create a system user and bind the user to the role of the tenant.
- The sub-tenant can further allocate the resources of its parent tenant. The sum of the resource percentages of direct sub-tenants under a parent tenant at each level cannot exceed 100%. The sum of the computing resource percentages of all level-1 tenants cannot exceed 100%.

Step 3 Check whether the current tenant needs to be associated with resources of other services.

- If yes, go to **Step 4**.
- If no, go to **Step 5**.

Step 4 Click **Associate Service** to configure other service resources used by the current tenant.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Set **Services** to **HBase**.
2. Set **Association Type** as follows:
   - **Exclusive** indicates that the service resources are used by the tenant exclusively and cannot be associated with other tenants.
   - **Shared** indicates that the service resources can be shared with other tenants.

NOTE

- Only HBase can be associated with a new tenant. However, HDFS, HBase, and Yarn can be associated with existing tenants.
- To associate an existing tenant with service resources, click the target tenant in the tenant list, switch to the **Service Associations** page, and click **Associate Service** to configure resources to be associated with the tenant.
- To disassociate an existing tenant from service resources, click the target tenant in the tenant list, switch to the **Service Associations** page, and click **Delete** in the **Operation** column. In the displayed dialog box, select **I have read the information and understand the impact** and click **OK**.

3. Click **OK**.

Step 5 Click **OK**. Wait until the system displays a message indicating that the tenant is successfully created.

----End

11.10.3 What Should I Do If a Job Fails to Be Submitted and the Error Is Related to OBS?

Question:

Why Does a Job Fail to Be Submitted and the Error Is Related to OBS?

Answer:

OBS errors are generally related to read and write. You can perform the following steps to locate the fault:

Step 1 View the failed job details on the **Jobs** page and check the actual job ID. If the job ID is empty, go to **Step 4**.

Step 2 Log in to FusionInsight Manager, choose **Cluster** > **Services** > **Yarn** > **Overview** > **ResourceManager Web UI**, and search for **logs** by the job ID.

Step 3 Check whether the following error information is displayed in the job log: If the following error message is displayed, bind an OBS agency to the cluster and submit the job again.



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 4 View the logs of to the Launcher job. On the **Jobs** page, view the logs to check whether the **stdout** or **stderr** log of the Launcher job contains error information. Locate the fault based on the error logs.

----End

11.10.4 Can I Run Multiple Spark Tasks at the Same Time After the Minimum Tenant Resources of an MRS Cluster Is Changed to 0?

Question:

Can I run multiple Spark tasks at the same time after the minimum tenant resources of an MRS cluster is changed to **0**?

Answer:

You can run only one Spark task at a time after the minimum tenant resources of an MRS cluster is changed to 0.

11.10.5 What Are the Differences Between the Client Mode and Cluster Mode of Spark Jobs?

You need to understand the concept ApplicationMaster before understanding the essential differences between Yarn-client and Yarn-cluster.

In Yarn, each application instance has an ApplicationMaster process, which is the first container started by the application. It interacts with ResourceManager and requests resources. After obtaining resources, it instructs NodeManager to start containers. The essential difference between the Yarn-cluster and Yarn-client modes lies in the ApplicationMaster process.

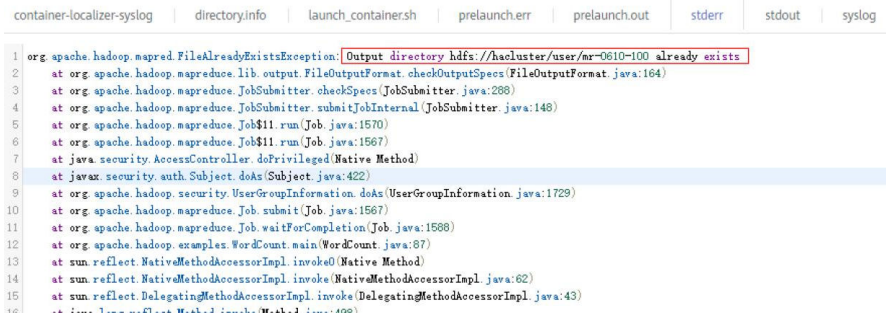
In Yarn-cluster mode, Driver runs in ApplicationMaster, which requests resources from Yarn and monitors the running status of a job. After a user submits a job, the client can be stopped and the job continues running on Yarn. Therefore, the Yarn-cluster mode is not suitable for running interactive jobs.

In Yarn-client mode, ApplicationMaster requests only Executor from Yarn. The client communicates with the requested containers to schedule tasks. Therefore, the client cannot be stopped.

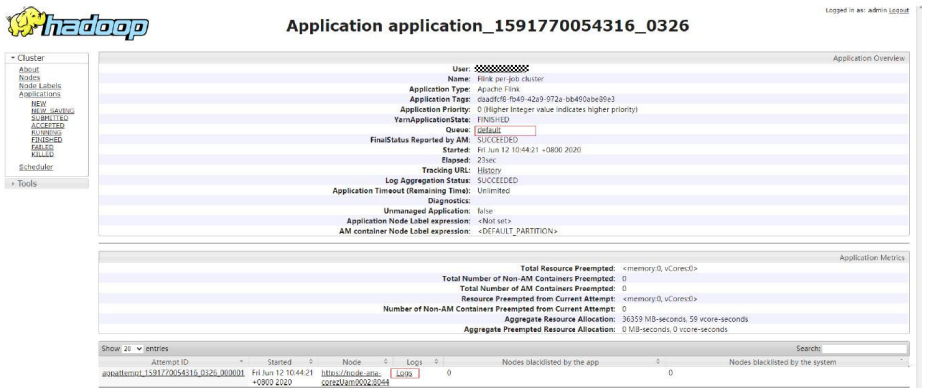
11.10.6 How Do I View MRS Job Logs?

Step 1 On the **Jobs** page of the MRS console, you can view logs of each job, including launcherJob and realJob logs.

- Generally, error logs are printed in **stderr** and **stdout** for launcherJob jobs, as shown in the following figure:



- You can view realJob logs on the ResourceManager web UI provided by the Yarn service on MRS Manager.



Step 2 Log in to the Master node of the cluster to obtain the job log files in **Step 1**. The HDFS path is **/tmp/logs/***{submit_user}*/**logs/***{application_id}*.

Step 3 After the job is submitted, if the job application ID cannot be found on the Yarn web UI, the job fails to be submitted. You can log in to the active Master node of the cluster and view the job submission process log **/var/log/executor/logs/exe.log**.

----End

11.10.7 What Can I Do If the System Displays a Message Indicating that the Current User Does Not Exist on Manager When I Submit a Job?

Question:

If IAM synchronization is not performed when a job is submitted to a cluster in security mode, the error message "The current user does not exist on MRS Manager. Grant the user sufficient permissions on IAM and then perform IAM user synchronization on the Dashboard tab page." is displayed.

Answer:

Before you submit the job, go to the **Dashboard** page, and click **Synchronize** on the right side of **IAM User Sync**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.8 What Can I Do If LauncherJob Fails to Be Executed and the Error Message "jobPropertiesMap is null" Is Displayed?

Question:

What should I do if the Launcher job fails because the user who submits the job does not have the write permission on the **hdfs /mrs/job-properties** directory?

Answer:

This problem has been fixed in the MRS 2.1.0.6. Additionally, you can also grant the write permission on the **//mrs/job-properties** directory to the synchronized user who submits the job on MRS Manager.

11.10.9 What Should I Do If the Flink Job Status on the MRS Console Is Inconsistent with That on YARN?

Question:

What should I do if the Flink job status on the MRS Console is inconsistent with that on YARN?

Answer:

To save storage space, you can modify the YARN configuration item **yarn.resourcemanager.max-completed-applications** to reduce the number of historical job records stored on YARN. Flink jobs are long-term jobs. The realJob is still running on Yarn, but the launcherJob has been deleted. As a result, the launcherJob cannot be found on Yarn, and the job status fails to be updated.

Stop the jobs where launcherJob cannot be found. The statuses of jobs submitted later will be updated.

11.10.10 What Should I Do If Error Code 403 Is Reported When a Spark Streaming Job Is Running?

Question:

What can I do if a Spark Streaming job fails after running for dozens of hours and error 403 is reported for OBS access?

Answer:

When a user submits a job that needs to read and write OBS, the job submission program adds the temporary access key (AK) and secret key (SK) for accessing OBS by default. However, the temporary AK and SK may expire.

If you want to run long-term jobs such as Flink and SparkStreaming jobs, you can enter the AK and SK in **Service Parameter** to prevent jobs execution failure due to key expiration.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.11 What Should I Do If Error Message "java.io.IOException: Connection reset by peer" Is Displayed During the Execution of a Spark Job?

Question:

The execution of a Spark job cannot finish and error message "java.io.IOException: Connection reset by peer" is displayed.

Answer:

Add the **executor.memoryOverhead** parameter to the job submission parameters.

11.10.12 What Should I Do If the Error Message "requestId=XXX" Is Displayed When a Spark Job Accesses OBS?

Question:

What do I do if the following error message is displayed when a Spark job accesses OBS?

```
xxx requestId=4971883851071737250 xxx
```

Answer:

1. Log in to the node where the Spark client is located, go to the **conf** directory, and change the value of the **fs.obs.metrics.switch** parameter in the **core-site.xml** configuration file to **false**.
2. Start the job again.

11.10.13 What Should I Do If the Error Message "UnknownScannerExeception" Is Displayed for Spark Jobs?

Symptom

Spark jobs run slowly. Warning information is printed in run logs, and the error cause is **UnknownScannerExeception**.

Solution

Before running a Spark job, adjust the value of **hbase.client.scanner.timeout.period** (for example, from 60 seconds to 120 seconds).

Log in to FusionInsight Manager and choose **Cluster** > **Services** > **HBase**. Click **Configurations** then **All Configurations**, search for **hbase.client.scanner.timeout.period**, and change its value to **120000** (unit: ms).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.14 What Can I Do If DataArts Studio Occasionally Fails to Schedule Spark Jobs?

Symptom

DataArtsStudio occasionally fails to schedule Spark jobs and the rescheduling also fails. The following error information is displayed:

```
Caused by: org.apache.spark.SparkException: Application application_1619511926396_2586346 finished with failed status
```

Solution

Log in to the node where the Spark client is located as user **root** and increase the value of the **spark.driver.memory** parameter in the **spark-defaults.conf** file.

11.10.15 What Should I Do If a Flink Job Fails to Execute and the Error Message "java.lang.NoSuchFieldError: SECURITY_SSL_ENCRYPT_ENABLED" Is Displayed?

Question:

A Flink job fails to be executed and the following error message is displayed:

```
Caused by: java.lang.NoSuchFieldError: SECURITY_SSL_ENCRYPT_ENABLED
```

Answer:

The third-party dependency package in the customer code conflicts with the cluster package. As a result, the job fails to be submitted to the MRS cluster.

You need to modify the dependency package, set the scope of the open source Hadoop package and Flink package in the POM file to **provide**, and pack and execute the job again.

11.10.16 What Should I Do If Submitted Yarn Jobs Cannot Be Viewed on the Web UI?

Question:

Why can't I view a created Yarn job when I log in to Manager as the **admin** user?

Answer:

- The **admin** user is a user on the cluster management page. Check whether the user has the **supergroup** permission. Generally, only the user with the **supergroup** permission can view jobs.
- Log in to Yarn as the user who submits the job. Do not use the **admin** user to view the job.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.17 What Can I Do If launcher-job Is Terminated by Yarn When a Flink Task Is Submitted?

Symptom

Launcher-job is terminated by Yarn due to insufficient heap size when a Flink task is submitted on the management plane.

Solution

Increase the heap size of the launcher-job queue.

1. Log in to the active OMS node as user **omm**.
2. Change the value of **job.launcher.resource.memory.mb** in **/opt/executor/webapps/executor/WEB-INF/classes/servicebroker.xml** to **2048**.
3. Run the **sh /opt/executor/bin/restart-executor.sh** command to restart the executor process.

11.10.18 What Should I Do If the Error Message "slot request timeout" Is Displayed When I Submit a Flink Job?

Symptom

When a Flink job is submitted, JobManager is started successfully. However, TaskManager remains in the starting state until timeout. The following error information is displayed:

```
org.apache.flink.runtime.jobmanager.scheduler.NoResourceAvailableException: Could not allocate the required slot within slot request timeout. Please make sure that the cluster has enough resources
```

Possible Causes

- The resources in the YARN queue are insufficient. As a result, TaskManager fails to start.
- Your JAR files conflict with those in the environment. You can execute the WordCount program to determine whether the issue occurs.
- If the cluster is in security mode, the SSL certificate of Flink may be incorrectly configured or has expired.

Solution

1. Add resources to the YARN queue.
2. Exclude the Flink and Hadoop dependencies in your JAR files so that Flink and Hadoop can depend only on the JAR files in the environment.
3. Reconfigure the SSL certificate of Flink..

11.10.19 FAQs About Importing and Exporting Data Using DistCP Jobs

- **Question**: Does a DistCP job check data consistency during data import and export?

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Answer: No. DistCP jobs only copy data but do not modify it.

- **Question**: When data is exported from a DistCP job, if some files already exist in OBS, how will the job process the files?

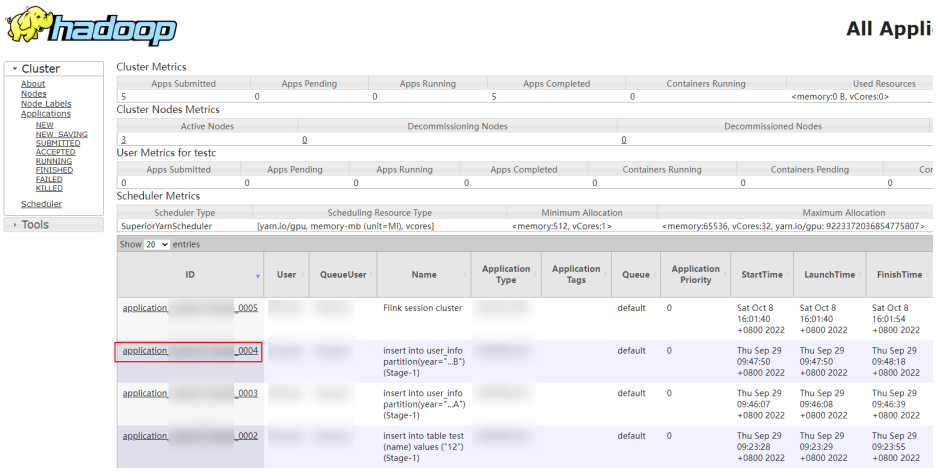
  Answer: DistCP jobs will overwrite the files in OBS.

11.10.20 How Do I View SQL Statements of Hive Jobs on the Yarn Web UI?

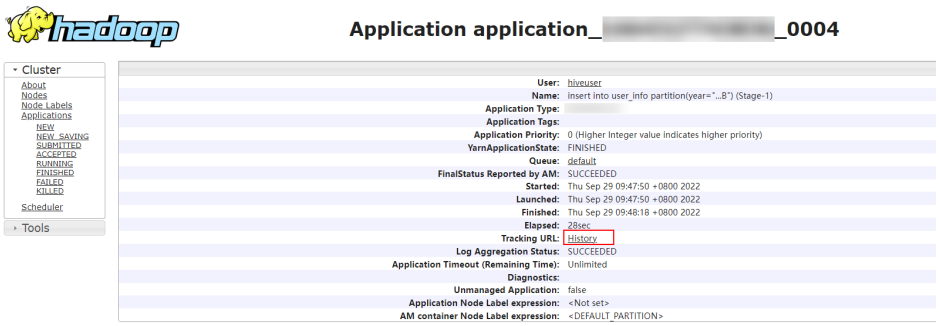
Step 1 Log in to FusionInsight Manager as a service user.

Step 2 Choose **Cluster** > **Services** > **Yarn**. Click **ResourceManager(xxx,Active)** next to **ResourceManager WebUI**.

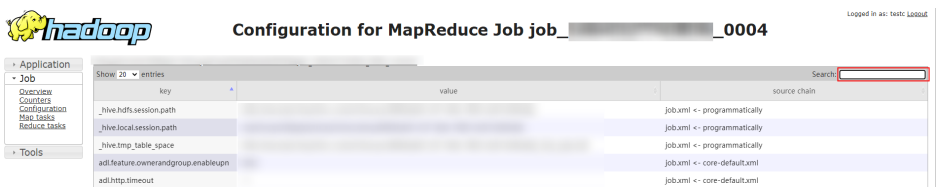
Step 3 On the YARN web UI, click the ID of the job to be viewed.



Step 4 Click **ApplicationMaster** or **History** next to **Tracking URL**.



Step 5 Choose **Configuration** in the left navigation pane and search for the **hive.query.string** parameter in the upper right corner to query the corresponding HiveSQL.



----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.10.21 How Do I View Logs of a Specified Yarn Task?

Step 1 Log in to the active Master node as user **root**.

Step 2 Run the following command to initialize environment variables:

source *Client installation directory*/**bigdata_env**

Step 3 If Kerberos authentication is enabled for the cluster, run the following command to authenticate the user. If Kerberos authentication is disabled, skip this step.

kinit *MRS cluster user*

Step 4 Run the following command to obtain the logs of a specified task:

yarn logs -applicationId *application_ID of the job you want to view*

----End

11.10.22 What Should I Do If a HiveSQL/HiveScript Job Fails to be Submitted After Hive Is Added?

This issue occurs because the **MRS CommonOperations** permission bound to the user group to which the user who submits the job belongs does not include the Hive permission after being synchronized to Manager. To solve this issue, perform the following operations:

1. Add the Hive service.
2. Log in to the IAM console and create a user group. The policy bound to the user group is the same as that of the user group to which the user who submits the job belongs.
3. Add the user who submits the job to the new user group.
4. Refresh the cluster details page on the MRS console. The status of IAM user synchronization is **Not synchronized**.
5. Click **Synchronize** on the right of **IAM User Sync**. Go back to the previous page. In the navigation pane on the left, choose **Operation Logs** and check whether the user is changed.
   - If yes, submit the Hive job again.
   - If no, check whether all the preceding operations are complete.
     - If yes, contact an O&M engineer.
     - If no, submit the Hive job after the preceding operations are complete.

11.10.23 Where Are the Execution Logs of Spark Jobs Stored?

Question:

Where Are the Execution Logs of Spark Jobs Stored?

Answer:

- Logs of unfinished Spark jobs are stored in the **/srv/BigData/hadoop/data1/nm/containerlogs/** directory on the Core node.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Logs of finished Spark jobs are stored in the **/tmp/logs/***Username***/logs** directory of the HDFS.

11.10.24 What Should I Do If an Alarm Indicating Insufficient Memory Is Reported During Spark Task Execution?

Symptom

When a Spark task is executed, an alarm indicating insufficient memory is reported. The alarm ID is 18022. As a result, no available memory can be used.

Procedure

Set the executor parameters in the SQL script to limit the number of cores and memory of an executor.

For example, the configuration is as follows:

```
set hive.execution.engine=spark;
set spark.executor.cores=2;
set spark.executor.memory=4G;
set spark.executor.instances=10;
```

Change the values of the parameters as required.

11.10.25 What Should I Do If It Takes a Long Time for Spark SQL to Access Hive Partitioned Tables Before a Job Starts?

Symptom

When Spark SQL is used to access Hive partitioned tables stored in OBS, the access speed is slow and a large number of OBS query APIs are called.

Example SQL:

```
select a,b,c from test where b=xxx
```

Fault Locating

According to the configuration, the task should scan only the partition whose b is *xxx*. However, the task logs show that the task scans all partitions and then calculates the data whose b is *xxx*. As a result, the task calculation is slow. In addition, a large number of OBS requests are sent because all files need to be scanned.

By default, the execution plan optimization based on partition statistics is enabled on MRS, which is equivalent to automatic execution of Analyze Table. (The default configuration method is to set **spark.sql.statistics.fallBackToHdfs** to **true**. You can set this parameter to **false**.) After this function is enabled, table partition statistics are scanned during SQL execution and used as cost estimation in the execution plan. For example, small tables identified during cost evaluation are broadcast to each node in the memory for join operations, significantly reducing shuffle time. This function greatly optimizes performance in join scenarios, but increases the number of OBS calls.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

Set the following parameter in Spark SQL and then run the SQL statement:

```
set spark.sql.statistics.fallBackToHdfs=false;
```

Alternatively, run the **--conf** command to set this parameter to **false** before startup.

```
--conf spark.sql.statistics.fallBackToHdfs=false
```

11.10.26 How Does the System Select the Queue When an MRS Cluster User Is Associated to Multiple Queues?

Question

Users A and B share the same user group and permissions in the MRS cluster. They are associated to both queue A and queue B.

Why are tasks of user A always submitted to queue A and tasks of user B always submitted to queue B?

Answer:

- If a user in the MRS cluster is associated to multiple queues, the system randomly selects a queue for tasks submitted by the user. You can set the queue permission of the user.and associate a user to only one queue.
- If you have to associate multiple queues, configure a global user policy to specify the default submission queue for a user.
  a. Take the MRS 3.x cluster as an example. Log in to Manager and click **Tenant Resources**.
  b. Choose **Dynamic Resource Plan** > **Global User Policy** > **Create Global User Policy**.

     Username: indicates the user for whom resource scheduling needs to be controlled. Enter an existing username in the current cluster.

     Default Queue: indicates the queue of the user. Enter the name of an existing queue in the current cluster.

11.11 Performance Tuning

11.11.1 How Do I Obtain the Hadoop Pressure Test Tool?

Question:

How Do I Obtain the Hadoop Pressure Test Tool?

Answer:

Obtain the Hadoop pressure test tool from the community at **https://github.com/Intel-bigdata/HiBench**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

11.11.2 How Do I Improve the Resource Utilization of Core Nodes in a Cluster?

1. Go to the Yarn service configuration page.
   - For versions earlier than MRS 1.8.10, log in to MRS Manager, choose **Services** > **Yarn** > **Service Configuration**, and select **All** from the **Basic** drop-down list.
   - For MRS 1.8.10 or later and MRS 2.*x*, click the cluster name. On the cluster details page that is displayed, click the **Components** tab and choose **Yarn**. Click the **Service Configuration** and select **All** from the **Basic** drop-down list.

     NOTE

     If the **Components** tab is unavailable, complete IAM user synchronization first. (On the **Dashboard** page, click **Synchronize** on the right side of **IAM User Sync** to synchronize IAM users.)
   - For MRS cluster version 3.*x* or later:

     Log in to FusionInsight Manager. Choose **Cluster** > *Name of the desired cluster* > **Services** > **Yarn** > **Configurations** > **All Configurations**.
2. Search for **yarn.nodemanager.resource.memory-mb**, and increase the value based on the actual memory of the cluster nodes.
3. Save the change and restart the affected services or instances.

11.11.3 How Do I Configure the Knox Memory for an MRS Cluster?

Step 1 Log in to a Master node of the cluster as user **root**.

Step 2 Run the following command on the Master node to open the **gateway.sh** file:

su omm

vim /opt/knox/bin/gateway.sh

Step 3 Change **APP_MEM_OPTS=""** to **APP_MEM_OPTS="-Xms256m -Xmx768m"**, save the file, and exit.

Step 4 Run the following command on the Master node to restart the knox process:

sh /opt/knox/bin/gateway.sh stop

sh /opt/knox/bin/gateway.sh start

Step 5 Repeat the preceding steps on each Master node.

Step 6 Run the **ps -ef |grep knox** command to check the configured memory.

Figure 11-9 knox memory



----End

11.11.4 How Do I Adjust the Memory Size of the manager-executor Process?

Symptom

The **manager-executor** process runs either on the Master1 or Master2 node in the MRS cluster in active/standby mode. This process is used to encapsulate the MRS management and control plane's operations on the MRS cluster, such as job submission, heartbeat reporting, certain alarm reporting, as well as cluster creation, scale-out, and scale-in. When you submit jobs on the MRS management and control plane, the Executor memory may become insufficient as the tasks increase or the number of concurrent tasks increases. As a result, the CPU usage is high and the Executor process experiences out-of-memory (OOM) errors.

Procedure

1. Log in to either the Master1 or Master2 node as user **root** and run the following command to switch to user **omm**:

   su - omm

2. Run the following command to modify the **catalina.sh** script. Specifically, search for **JAVA_OPTS** in the script, find the configuration items similar to **JAVA_OPTS="-Xms1024m -Xmx4096m**, and change the values of the items to desired ones, and save the modification.

   vim /opt/executor/bin/catalina.sh

3. The **manager-executor** process only runs on either the Master1 or Master2 node in active/standby mode. Check whether it exists on the node before restarting it.

   a. Log in to the Master1 and Master2 nodes and run the following command to check whether the process exists. If any command output is displayed, the process exists.

      ps -ef | grep "/opt/executor" | grep -v grep

   b. Run the following command to restart the process:

      sh /opt/executor/bin/shutdown.sh

      sh /opt/executor/bin/startup.sh

11.11.5 What Should I Do If the spark.yarn.executor.memoryOverhead Setting Does Not Take Effect?

Symptom

The overhead memory of the executor needs to be adjusted for Spark tasks. The **spark.yarn.executor.memoryOverhead** parameter is set to **4096**. However, the default value **1024** is used to apply for resources during actual computation.

Fault Locating

In Spark 2.3 and later versions, use the new parameter **spark.executor.memoryOverhead** to set the overhead memory of the executor. If

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

both old and new parameters are set, the value of **spark.yarn.executor.memoryOverhead** does not take effect, and the value of **spark.executor.memoryOverhead** is used.

Same thing happens if you use **spark.driver.memoryOverhead** to set the overhead memory of driver.

Procedure

Use the new parameter:

```
spark.executor.memoryOverhead=4096
```

11.12 Application Development

11.12.1 How Do I Get My Data into OBS or HDFS?

MRS can process data in OBS and HDFS. You can get your data into OBS or HDFS as follows:

1. Upload local data to OBS.
   a. Log in to the OBS console.
   b. Create a parallel file system named **userdata** on OBS and create the **program**, **input**, **output**, and **log** folders in the file system.
      i. Choose **Parallel File System** > **Create Parallel File System**, and create a file system named **userdata**.
      ii. In the OBS file system list, click the file system name **userdata**, choose **Files** > **Create Folder**, and create the **program**, **input**, **output**, and **log** folders.
   c. Upload data to the **userdata** file system.
      i. Go to the **program** folder and click **Upload File**.
      ii. Click **add file** and select a user program.
      iii. Click **Upload**.
      iv. Upload the user data file to the **input** directory using the same method.
2. Import OBS data to HDFS.

   You can import OBS data to HDFS only when **Kerberos Authentication** is disabled and the cluster is running.
   a. Log in to the MRS console.
   b. Click the name of the cluster.
   c. On the page displayed, select the **Files** tab page and click **HDFS File List**.
   d. Select a data directory, for example, **bd_app1**.

      The **bd_app1** directory is only an example. You can use any directory on the page or create a new one.
   e. Click **Import Data** and click **Browse** to select an OBS path and an HDFS path.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

f. Click **OK**.

You can view the file upload progress on the **File Operation Records** tab page.

11.12.2 Can MRS Write Data to HBase Through an HBase External Table of Hive?

Question:

Can MRS write data to HBase through an HBase external table of Hive?

Answer:

This operation is not supported.

Hive on HBase supports only data query.

11.12.3 Does OpenTSDB Support Python APIs?

Question:

Does OpenTSDB support Python APIs?

Answer:

OpenTSDB supports Python APIs. OpenTSDB provides HTTP-based RESTful APIs that are language-independent. Any language that supports HTTP requests can interconnect to OpenTSDB.

11.12.4 How Do I Configure the node_id Parameter When Calling the API for Adjusting Cluster Nodes?

Question:

How Do I Configure the Request Parameter node_id When Calling the API for Adjusting Cluster Nodes (/v1.1/{project_id}/cluster_infos/{cluster_id})?

Answer:

When the API is called, the value of the request parameter **node_id** is fixed to **node_orderadd**.

11.12.5 How Do I Manage and Use Third-Party JAR Packages for MRS Cluster Components?

Some components within an MRS cluster support the use of custom third-party JAR packages to meet specific customization needs. When a user application invokes a third-party JAR package while using a component, it can be stored in a designated directory on each node. Upon restarting the component, the JAR package in that directory will be automatically loaded by the corresponding component program.

The directory is as follows: *<Cluster installation directory>***/third_lib/***<Component name>* of the component installation node.

Here are two examples of saving JAR files to their respective directories:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Save the JAR package file used by the HBase service to the **${BIGDATA_HOME}/third_lib/HBase** directory of the HBase node.
- Save the JAR package file used by the Hive service to the **${BIGDATA_HOME}/third_lib/Hive** directory of each HiveServer instance node.

NOTE

- Currently, custom JAR files are supported on HDFS, YARN, MapReduce, Spark, HBase, and Hive.
- Hive supports only Hive custom JAR files developed by users for UDF implementation. Do not store other open-source JAR files in the directory, including **org.apache.***, **com.google.***, and **javax.***.
- This directory is supported only by MRS 3.x clusters.

11.13 Upgrade and Patching

11.13.1 How Do I Upgrade an MRS Cluster?

- Non-LTS clusters

  An existing non-LTS cluster cannot be upgraded to a later version. You can create a cluster of the new version and migrate data from the old cluster to the new cluster.

- MRS LTS clusters

  MRS LTS clusters can be upgraded to a target LTS version. Contact our O&M engineers to confirm the upgrade path and apply for professional services for cluster upgrade.

11.13.2 Does MRS Support Kernel Version Upgrade of Components in a Cluster?

Question:

Does MRS support kernel version upgrade of components in a cluster?

Answer:

MRS does not support independent upgrade of kernel versions of components in a cluster. To use components of new versions, delete the current cluster and create another cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.