

**MapReduce Service**

# Troubleshooting

**Issue** 01  
**Date** 2024-12-18



**Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2024. 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**



HUAWEI 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.

---

# Contents

---

<b>1 Account Passwords.....</b>	<b>1</b>
1.1 Resetting or Changing the Password of Manager User admin.....	1
1.2 Failed to Download Authentication Credentials If the Username Is Too Long.....	1
<b>2 Account Permissions.....</b>	<b>4</b>
2.1 A Message Is Displayed Indicating That the User Does Not Have the Permission to Obtain the MRS Cluster Hosts.....	4
2.2 Failed to View MRS Cluster Details.....	5
<b>3 Common Exceptions in Logging In to the Cluster Manager.....</b>	<b>6</b>
3.1 Failed to Access Manager of an MRS Cluster.....	6
3.2 Accessing the Web Pages.....	8
3.2.1 Error "502 Bad Gateway" Is Reported During the Access to MRS Manager.....	8
3.2.2 An Error Message Occurs Indicating that the VPC Request Is Incorrect During the Access.....	9
3.2.3 Error 503 Is Reported When Manager Is Accessed Through Direct Connect.....	10
3.2.4 Error Message "You have no right to access this page." Is Displayed When Users log in to the Cluster Page.....	11
3.2.5 Error Message "Invalid credentials" Is Displayed When a User Logs In to Manager.....	12
3.2.6 Failed to Log In to the Manager After Timeout.....	13
3.2.7 Failed to Log In to MRS Manager After the Python Upgrade.....	13
3.2.8 Failed to Log In to MRS Manager After Changing the Domain Name.....	14
3.2.9 Manager Page Is Blank After a Success Login.....	15
3.2.10 Cluster Login Fails Because Native Kerberos Is Installed on Cluster Nodes.....	16
3.2.11 Using Google Chrome to Access MRS Manager on macOS.....	17
3.2.12 How Do I Unlock a User Who Logs in to Manager?.....	18
3.2.13 Why Does the Manager Page Freeze?.....	19
3.3 Common Exceptions in Accessing the MRS Web UI.....	20
3.3.1 How Do I Do If an Error Is Reported or Some Functions Are Unavailable When I Access the Web UIs of HDFS, Hue, YARN, HetuEngine, and Flink?.....	20
3.3.2 Error 500 Is Reported When a User Accesses the Component Web UI.....	21
3.3.3 [HBase WebUI] Users cannot switch from the HBase WebUI to the RegionServer WebUI.....	22
3.3.4 [HDFS WebUI] When users access the HDFS WebUI, an error message is displayed indicating that the number of redirections is too large.....	23
3.3.5 [HDFS WebUI] Failed to access the HDFS WebUI using the Internet Explorer.....	23
3.3.6 [Hue Web UI] A "No Permission" Error Is Displayed When a User Log In to the Hue Web UI.....	24

3.3.7 [Hue Web UI] Failed to Access the Hue Web UI.....	25
3.3.8 [Hue WebUI] The error "Proxy Error" is reported when a user accesses the Hue WebUI.....	26
3.3.9 [Hue WebUI] Why Is the Hue Native Page Cannot Be Properly Displayed If the Hive Service Is Not Installed in a Cluster?.....	27
3.3.10 Hue (Active) Cannot Open Web Pages.....	27
3.3.11 [Ranger WebUI] Why Cannot a New User Log In to Ranger After Changing the Password?.....	28
3.3.12 [Tez WebUI] Error 404 is reported when users access the Tez WebUI.....	28
3.3.13 [Spark WebUI] Why Cannot I Switch from the Yarn Web UI to the Spark Web UI?.....	29
3.3.14 [Spark WebUI] What Can I Do If an Error Occurs when I Access the Application Page Because the Application Cached by HistoryServer Is Recycled?.....	30
3.3.15 [Spark WebUI] Why Is the Native Page of an Application in Spark2x JobHistory Displayed Incorrectly?.....	31
3.3.16 [Spark WebUI] The Spark2x WebUI fails to be accessed using the Internet Explorer.....	32
3.3.17 [Yarn Web UI] Failed to Access the Yarn Web UI.....	32
<b>4 APIs.....</b>	<b>34</b>
4.1 Failed to Call an API to Create a Cluster.....	34
<b>5 Cluster Management.....</b>	<b>35</b>
5.1 Failed to Reduce Task Nodes.....	35
5.2 OBS Certificate in a Cluster Expired.....	36
5.3 Replacing a Disk in an MRS Cluster (Applicable to 2.x and Earlier).....	38
5.4 Replacing a Disk in an MRS Cluster (Applicable to 3.x).....	40
5.5 Failed to Execute an MRS Backup Task.....	43
5.6 Inconsistency Between df and du Command Output on the Core Node.....	44
5.7 Disassociating a Subnet from a Network ACL.....	45
5.8 MRS Cluster Becomes Abnormal After the Hostname of a Node Is Changed.....	45
5.9 Processes Are Terminated Unexpectedly.....	46
5.10 Failed to Configure Cross-Cluster Mutual Trust for MRS.....	48
5.11 Network Is Unreachable When Python Is Installed on an MRS Cluster Node Using Pip3.....	50
5.12 Connecting the Open-Source confluent-kafka-go to an MRS Security Cluster.....	50
5.13 Failed to Execute the Periodic Backup Task of an MRS Cluster.....	52
5.14 Failed to Download the MRS Cluster Client.....	53
5.15 An Error Is Reported When a Flink Job Is Submitted in an MRS Cluster with Kerberos Authentication Enabled.....	53
5.16 An Error Is Reported When the Insert Command Is Executed on the Hive Beeline CLI.....	56
5.17 Upgrading the OS to Fix Vulnerabilities for an MRS Cluster Node.....	56
5.18 Failed to Migrate Data to MRS HDFS Using CDM.....	58
5.19 Alarms Indicating Heartbeat Interruptions Between Nodes Are Frequently Generated in the MRS Cluster.....	59
5.20 High Memory Usage of the PMS Process.....	61
5.21 High Memory Usage of the Knox Process.....	62
5.22 It Takes a Long Time to Access HBase from a Client Outside a Security Cluster.....	63
5.23 Failed to Submit Jobs.....	64
5.24 OS Disk Space Is Insufficient Due to Oversized HBase Log Files.....	67

5.25 OS Disk Space Is Insufficient Due to Oversized HDFS Log Files.....	67
5.26 An Exception Occurs During Specifications Upgrade of Nodes in an MRS Cluster.....	68
5.27 Failed to Delete a New Tenant on FusionInsight Manager.....	70
5.28 MRS Cluster Becomes Unavailable After the VPC Is Changed.....	71
5.29 Failed to Submit Jobs on the MRS Console.....	73
5.30 Error "symbol xxx not defined in file libcrypto.so.1.1" Is Displayed During HA Certificate Generation .....	74
5.31 Some Instances Fail to Be Started After Core Nodes Are Added to the MRS Cluster.....	75
<b>6 Using Alluixio.....</b>	<b>78</b>
6.1 Error Message "Does not contain a valid host:port authority" Is Reported When Alluixio Is in HA Mode .....	78
<b>7 Using ClickHouse.....</b>	<b>79</b>
7.1 ClickHouse Fails to Start Due to Incorrect Data in ZooKeeper.....	79
7.2 An Exception Occurs When ClickHouse Consumes Kafka Data.....	81
<b>8 Using DBService.....</b>	<b>83</b>
8.1 DBServer Instance Is in Abnormal Status.....	83
8.2 DBServer Instance Remains in the Restoring State.....	84
8.3 Default Port 20050 or 20051 of DBService Is Occupied.....	85
8.4 DBServer Instance Is Always in the Restoring State Because the Incorrect /tmp Directory Permission .....	86
8.5 Failed to Execute a DBService Backup Task.....	87
8.6 Components Failed to Connect to DBService in Normal State.....	88
8.7 DBServer Failed to Start.....	88
8.8 DBService Backup Failed Because the Floating IP Address Is Unreachable.....	89
8.9 DBService Failed to Start Due to the Loss of the DBService Configuration File.....	91
<b>9 Using Flink.....</b>	<b>94</b>
9.1 Error Message "Error While Parsing YAML Configuration File: Security.kerberos.login.keytab" Is Displayed When a Command Is Executed on the Flink Client.....	94
9.2 Error Message "Error while parsing YAML configuration file : security.kerberos.login.principal:pippo" Is Displayed When a Command Is Executed on the Flink Client.....	95
9.3 Error Message "Could Not Connect to the Leading JobManager" Is Displayed When a Command Is Executed on the Flink Client.....	96
9.4 Failed to Create a Flink Cluster by Running yarn-session As Different Users.....	97
9.5 Flink Service Program Fails to Read Files on the NFS Disk.....	98
9.6 Failed to Customize the Flink Log4j Log Level.....	99
<b>10 Using Flume.....</b>	<b>101</b>
10.1 Class Cannot Be Found After Flume Submits Jobs to Spark Streaming.....	101
10.2 Failed to Install a Flume Client.....	102
10.3 A Flume Client Cannot Connect to the Server.....	102
10.4 Flume Data Fails to Be Written to the Component.....	103
10.5 Flume Server Process Fault.....	104
10.6 Flume Data Collection Is Slow.....	104

10.7 Failed to Start Flume.....	105
<b>11 Using HBase.....</b>	<b>107</b>
11.1 Slow Response to HBase Connection.....	107
11.2 Failed to Authenticate the HBase User.....	108
11.3 RegionServer Failed to Start Because the Port Is Occupied.....	108
11.4 HBase Failed to Start Due to Insufficient Node Memory.....	109
11.5 HBase Service Unavailable Due to Poor HDFS Performance.....	109
11.6 HBase Failed to Start Due to Inappropriate Parameter Settings.....	111
11.7 RegionServer Failed to Start Due to Residual Processes.....	111
11.8 HBase Failed to Start Due to a Quota Set on HDFS.....	112
11.9 HBase Failed to Start Due to Corrupted Version Files.....	113
11.10 High CPU Usage Caused by Zero-Loaded RegionServer.....	113
11.11 HBase Failed to Start with "FileNotFoundException" in RegionServer Logs.....	115
11.12 The Number of RegionServers Displayed on the Native Page Is Greater Than the Actual Number After HBase Is Started.....	116
11.13 RegionServer Instance Is in the Restoring State.....	117
11.14 HBase Failed to Start in a Newly Installed Cluster.....	118
11.15 HBase Failed to Start Due to the Loss of the ACL Table Directory.....	119
11.16 HBase Failed to Start After the Cluster Is Powered Off and On.....	120
11.17 Failed to Import HBase Data Due to Oversized File Blocks.....	121
11.18 Failed to Load Data to the Index Table After an HBase Table Is Created Using Phoenix.....	122
11.19 Failed to Run the <b>hbase shell</b> Command on the MRS Cluster Client.....	124
11.20 Disordered Information Display on the HBase Shell Client Console Due to Printing of the INFO Information.....	124
11.21 HBase Failed to Start Due to Insufficient RegionServer Memory.....	126
11.22 Failed to Start HRegionServer on the Node Newly Added to the Cluster.....	126
11.23 Region in the RIT State for a Long Time Due to HBase File Loss.....	127
<b>12 Using HDFS.....</b>	<b>129</b>
12.1 HDFS NameNode Instances Become Standby After the RPC Port Is Changed.....	129
12.2 An Error Is Reported When the HDFS Client Is Connected Through a Public IP Address.....	130
12.3 Failed to Use Python to Remotely Connect to the Port of HDFS.....	131
12.4 HDFS Capacity Reaches 100%, Causing Unavailable Upper-Layer Services Such as HBase and Spark.....	132
12.5 Error Message "Permission denied" Is Displayed When HDFS and Yarn Are Started.....	133
12.6 HDFS Users Can Create or Delete Files in Directories of Other Users.....	134
12.7 A DataNode of HDFS Is Always in the Decommissioning State.....	135
12.8 HDFS NameNode Failed to Start Due to Insufficient Memory.....	137
12.9 A Large Number of Blocks Are Lost in HDFS due to the Time Change Using ntpdate.....	139
12.10 CPU Usage of DataNodes Is Close to 100% Occasionally, Causing Node Loss.....	141
12.11 Manually Performing Checkpoints When a NameNode Is Faulty for a Long Time.....	142
12.12 Error "Failed to place enough replicas" Is Reported When HDFS Reads or Writes Files.....	143
12.13 Maximum Number of File Handles Is Set to a Too Small Value, Causing File Reading and Writing Exceptions.....	144

12.14 HDFS Client File Fails to Be Closed After Data Writing.....	145
12.15 File Fails to Be Uploaded to HDFS Due to File Errors.....	147
12.16 After dfs.blocksize Is Configured on the UI and Data Is Uploaded, the Block Size Does Not Change .....	148
12.17 HDFS File Fails to Be Read, and Error Message "FileNotFoundException" Is Displayed.....	149
12.18 Failed to Write Files to HDFS, and Error Message "item limit of xxx is exceeded" Is Displayed.....	150
12.19 Adjusting the Log Level of the HDFS SHDFSshell Client.....	150
12.20 HDFS File Read Fails, and Error Message "No common protection layer" Is Displayed.....	151
12.21 Failed to Write Files Because the HDFS Directory Quota Is Insufficient.....	152
12.22 Balancing Fails, and Error Message "Source and target differ in block-size" Is Displayed.....	153
12.23 Failed to Query or Delete HDFS Files.....	154
12.24 Uneven Data Distribution Due to Non-HDFS Data Residuals.....	155
12.25 Uneven Data Distribution Due to HDFS Client Installation on the DataNode.....	156
12.26 Unbalanced DataNode Disk Usages of a Node.....	156
12.27 Locating Common Balance Problems.....	157
12.28 HDFS Displays Insufficient Disk Space But 10% Disk Space Remains.....	158
12.29 Error Message "error creating DomainSocket" Is Displayed When the HDFS Client Installed on the Core Node in a Normal Cluster Is Used.....	159
12.30 HDFS Files Fail to Be Uploaded When the Client Is Installed on a Node Outside the Cluster.....	159
12.31 Insufficient Number of Replicas Is Reported During High Concurrent HDFS Writes.....	160
12.32 HDFS Client Failed to Delete Overlong Directories.....	161
12.33 An Error Is Reported When a Node Outside the Cluster Accesses MRS HDFS.....	163
12.34 "ALM-12027 Host PID Usage Exceeds the Threshold" Is Generated for a NameNode.....	164
12.35 ALM-14012 JournalNode Is Out of Synchronization Is Generated in the Cluster.....	166
12.36 Failed to Decommission a DataNode Due to HDFS Block Loss.....	166
12.37 An Error Is Reported When DistCP Is Used to Copy an Empty Folder.....	168
<b>13 Using Hive.....</b>	<b>169</b>
13.1 Common Hive Logs.....	169
13.2 Failed to Start Hive.....	170
13.3 Error Message "Cannot modify xxx at runtime" Is Displayed When the set Command Is Executed in a Security Cluster.....	171
13.4 Specifying a Queue When Submitting a Hive Task.....	172
13.5 Setting the Map/Reduce Memory on the Client.....	173
13.6 Specifying the Output File Compression Format When Importing a Hive Table.....	173
13.7 Description of the Hive Table Is Too Long to Be Completely Displayed.....	174
13.8 NULL Is Displayed When Data Is Inserted After the Partition Column Is Added to a Hive Table.....	175
13.9 New User Created in the Cluster Does Not Have the Permission to Query Hive Data.....	175
13.10 An Error Is Reported When SQL Is Executed to Submit a Task to a Specified Queue.....	177
13.11 An Error Is Reported When the "load data inpath" Command Is Executed.....	177
13.12 An Error Is Reported When the "load data local inpath" Command Is Executed.....	178
13.13 An Error Is Reported When the create external table Command Is Executed.....	179
13.14 An Error Is Reported When the dfs -put Command Is Executed on the Beeline Client.....	180
13.15 Insufficient Permissions to Execute the set role admin Command.....	180

13.16 An Error Is Reported When a UDF Is Created on the Beeline Client.....	181
13.17 Hive Is Faulty.....	182
13.18 Difference Between Hive Service Health Status and Hive Instance Health Status.....	182
13.19 "authentication failed" Is Displayed During an Attempt to Connect to the Shell Client.....	183
13.20 Failed to Access ZooKeeper from the Client.....	183
13.21 "Invalid function" Is Displayed When a UDF Is Used.....	185
13.22 Hive Service Status Is Unknown.....	185
13.23 Health Status of a HiveServer or MetaStore Instance Is unknown.....	186
13.24 Health Status of a HiveServer or MetaStore Instance Is Concerning.....	186
13.25 Garbled Characters Returned Upon a Query If Text Files Are Compressed Using ARC4.....	187
13.26 Hive Task Failed to Run on the Client but Successful on Yarn.....	187
13.27 Error Message "Execution Error return code 2" Is Displayed When the SELECT Statement Is Executed.....	188
13.28 Failed to Perform drop partition When There Are a Large Number of Partitions.....	190
13.29 Failed to Start the Local Task When the Join Operation Is Performed.....	190
13.30 WebHCat Fails to Be Started After the Hostname Is Changed.....	192
13.31 An Error Is Reported When the Hive Sample Program Is Running After the Domain Name of a Cluster Is Changed.....	193
13.32 Hive MetaStore Exception Occurs When the Number of DBService Connections Exceeds the Upper Limit.....	194
13.33 Error Message "Failed to execute session hooks: over max connections" Is Displayed on the Beeline Client.....	195
13.34 Error Message "OutOfMemoryError" Is Displayed on the Beeline Client.....	196
13.35 Task Execution Fails Because the Input File Number Exceeds the Threshold.....	198
13.36 Hive Task Execution Fails Because of Stack Memory Overflow.....	199
13.37 Task Failed Due to Concurrent Writes to One Table or Partition.....	200
13.38 Hive Task Failed Due to a Lack of HDFS Directory Permission.....	201
13.39 Failed to Load Data to Hive Tables.....	202
13.40 Failed to Run the Application Developed Based on the Hive JDBC Code Case.....	203
13.41 HiveServer and HiveHCat Process Faults.....	204
13.42 Error Message "ConnectionLoss for hiveserver2" Is Displayed When MRS Hive Connects to ZooKeeper.....	205
13.43 An Error Is Reported When Hive Executes the insert into Statement.....	206
13.44 Timeout Reported When Adding the Hive Table Field.....	207
13.45 Failed to Restart Hive.....	210
13.46 Failed to Delete a Table Due to Excessive Hive Partitions.....	211
13.47 An Error Is Reported When msck repair table Is Executed on Hive.....	212
13.48 Insufficient User Permission for Running the insert into Command on Hive.....	212
13.49 Releasing Disk Space After Dropping a Table in Hive.....	213
13.50 Abnormal Hive Query Due to Damaged Data in the JSON Table.....	214
13.51 Connection Timed Out During SQL Statement Execution on the Hive Client.....	216
13.52 WebHCat Failed to Start Due to Abnormal Health Status.....	217
13.53 WebHCat Failed to Start Because the mapred-default.xml File Cannot Be Parsed.....	218



13.54 An SQL Error Is Reported When the Number of MetaStore Dynamic Partitions Exceeds the Threshold.....	219
<b>14 Using Hue.....</b>	<b>220</b>
14.1 An Unknown Job Is Running on the Hue Page.....	220
14.2 HQL Fails to Be Executed on Hue Using Internet Explorer.....	221
14.3 Failed to Access the Hue Web UI.....	221
14.4 HBase Tables Cannot Be Loaded on the Hue Web UI.....	222
<b>15 Using Impala.....</b>	<b>224</b>
15.1 Failed to Connect to impala-shell.....	224
15.2 Failed to Create a Kudu Table.....	225
15.3 Installing Python2 on the Impala Client.....	225
<b>16 Using Kafka.....</b>	<b>229</b>
16.1 An Error Is Reported When the Kafka Client Is Run to Obtain Topics.....	229
16.2 Using Python3.x to Connect to Kafka in a Security Cluster.....	229
16.3 Flume Normally Connects to Kafka but Fails to Send Messages.....	232
16.4 Producer Fails to Send Data and Error Message "NullPointerException" Is Displayed.....	233
16.5 Producer Fails to Send Data and Error Message "TOPIC_AUTHORIZATION_FAILED" Is Displayed....	235
16.6 Producer Occasionally Fails to Send Data and the Log Displays "Too many open files in system" ...	238
16.7 Consumer Is Initialized Successfully, but the Specified Topic Message Cannot Be Obtained from Kafka.....	240
16.8 Consumer Fails to Consume Data and Remains in the Waiting State.....	244
16.9 SparkStreaming Fails to Consume Kafka Messages, and "Error getting partition metadata" Is Displayed.....	246
16.10 Consumer Fails to Consume Data in a Newly Created Cluster, and Message "GROUP_COORDINATOR_NOT_AVAILABLE" Is Displayed.....	248
16.11 SparkStreaming Fails to Consume Kafka Messages, and Message "Couldn't find leader offsets" Is Displayed.....	249
16.12 Consumer Fails to Consume Data and Message "SchemaException: Error reading field" Is Displayed.....	251
16.13 Kafka Consumer Loses Consumed Data.....	252
16.14 Failed to Start Kafka Due to Account Lockout.....	253
16.15 Kafka Broker Reports Abnormal Processes and the Log Shows "IllegalArgumentExcepion".....	253
16.16 Kafka Topics Cannot Be Deleted.....	255
16.17 Error "AdminOperationException" Is Displayed When a Kafka Topic Is Deleted.....	259
16.18 When a Kafka Topic Fails to Be Created, "NoAuthExcepion" Is Displayed.....	261
16.19 Failed to Set an ACL for a Kafka Topic, and "NoAuthExcepion" Is Displayed.....	263
16.20 When a Kafka Topic Fails to Be Created, "NoNode for /brokers/ids" Is Displayed.....	265
16.21 When a Kafka Topic Fails to Be Created, "replication factor larger than available brokers" Is Displayed.....	266
16.22 Consumer Repeatedly Consumes Data.....	267
16.23 Leader for the Created Kafka Topic Partition Is Displayed as none.....	269
16.24 Safety Instructions on Using Kafka.....	271
16.25 Obtaining Kafka Consumer Offset Information.....	276

16.26 Adding or Deleting Configurations for a Topic.....	278
16.27 Reading the Content of the __consumer_offsets Internal Topic.....	279
16.28 Configuring Logs for Shell Commands on the Kafka Client.....	280
16.29 Obtaining Topic Distribution Information.....	281
16.30 Kafka HA Usage Description.....	283
16.31 Failed to Manage a Kafka Cluster Using the Kafka Shell Command.....	286
16.32 Kafka Producer Writes Oversized Records.....	287
16.33 Kafka Consumer Reads Oversized Records.....	288
16.34 High Usage of Multiple Disks on a Kafka Cluster Node.....	289
16.35 Kafka Is Disconnected from the ZooKeeper Client.....	291
<b>17 Using Oozie.....</b>	<b>293</b>
17.1 Oozie Jobs Do Not Run When a Large Number of Jobs Are Submitted Concurrently.....	293
17.2 An Error Is Reported When Oozie Schedules HiveSQL Jobs.....	294
17.3 Oozie Tasks Cannot Be Submitted on a Client Outside the MRS Cluster or Can Be Submitted Only Two Hours Later.....	294
<b>18 Using Presto.....</b>	<b>296</b>
18.1 During sql-standard-with-group Configuration, a Schema Fails to Be Created and the Error Message "Access Denied" Is Displayed.....	296
18.2 Coordinator Process of Presto Cannot Be Started.....	297
18.3 When Presto Queries a Kudu Table, an Error Is Reported Indicating That the Table Cannot Be Found.....	299
18.4 No Data is Found in the Hive Table Using Presto.....	300
18.5 Error Message "The node may have crashed or be under too much load" Is Displayed During MRS Presto Query.....	301
18.6 Accessing Presto from an MRS Cluster Through a Public Network.....	302
<b>19 Using Spark.....</b>	<b>304</b>
19.1 An Error Is Reported When the Split Size Is Changed for a Running Spark Application.....	304
19.2 Incorrect Parameter Format Is Displayed When a Spark Task Is Submitted.....	305
19.3 Spark, Hive, and Yarn Are Unavailable Due to Insufficient Disk Capacity.....	306
19.4 A Spark Job Fails to Run Due to Incorrect JAR File Import.....	307
19.5 Spark Job Suspended Due to Insufficient Memory or Lack of JAR Packages.....	307
19.6 Error "ClassNotFoundException" Is Reported When a Spark Task Is Submitted.....	309
19.7 Driver Displays a Message Indicating That the Running Memory Exceeds the Threshold When a Spark Task Is Submitted.....	310
19.8 Error "Can't get the Kerberos realm" Is Reported When a Spark Task Is Submitted in Yarn-Cluster Mode.....	311
19.9 Failed to Start spark-sql and spark-shell Due to JDK Version Mismatch.....	312
19.10 ApplicationMaster Fails to Start Twice When a Spark Task Is Submitted in Yarn-client Mode.....	313
19.11 Failed to Connect to ResourceManager When a Spark Task Is Submitted.....	314
19.12 DataArts Studio Failed to Schedule Spark Jobs.....	315
19.13 Job Status Is error After a Spark Job Is Submitted Through an API.....	316
19.14 ALM-43006 Is Repeatedly Reported for the MRS Cluster.....	317
19.15 Failed to Create or Delete a Table in Spark Beeline.....	317

19.16 Failed to Connect to the Driver When a Spark Job Is Submitted on a Node Outside the Cluster....	319
19.17 Large Number of Shuffle Results Are Lost During Spark Task Execution.....	320
19.18 Disk Space Is Insufficient Due to Long-Term Running of JDBCServer.....	321
19.19 Failed to Load Data to a Hive Table Across File Systems by Running SQL Statements Using Spark Shell.....	322
19.20 Spark Task Submission Failure.....	323
19.21 Spark Task Execution Failure.....	324
19.22 JDBCServer Connection Failure.....	324
19.23 Failed to View Spark Task Logs.....	325
19.24 Spark Streaming Task Submission Issues.....	326
19.25 Authentication Fails When Spark Connects to Other Services.....	326
19.26 Authentication Fails When Spark Connects to Kafka.....	327
19.27 An Error Occurs When SparkSQL Reads the ORC Table.....	328
19.28 Failed to Switch to the Log Page from stderr and stdout on the Native Spark Web UI.....	328
19.29 An Error Is Reported When spark-beeline Is Used to Query a Hive View.....	331
<b>20 Using Sqoop.....</b>	<b>333</b>
20.1 Connecting Sqoop to MySQL.....	333
20.2 Failed to Find the <b>HBaseAdmin.&lt;init&gt;</b> Method When Sqoop Reads Data from the MySQL Database to HBase.....	334
20.3 An Error Is Reported When a Sqoop Task Is Created Using Hue to Import Data from HBase to HDFS.....	335
20.4 A Data Format Error Is Reported When Data Is Exported from Hive to MySQL 8.0 Using Sqoop....	339
20.5 An Error Is Reported When the sqoop import Command Is Executed to Extract Data from PostgreSQL to Hive.....	341
20.6 Failed to Use Sqoop to Read MySQL Data and Write Parquet Files to OBS.....	342
20.7 An Error Is Reported When Database Data Is Migrated Using Sqoop.....	343
<b>21 Using Storm.....</b>	<b>345</b>
21.1 Invalid Hyperlink of Events on the Storm Web UI.....	345
21.2 Failed to Submit the Storm Topology.....	346
21.3 Failed to Submit the Storm Topology and Message "Failed to check principle for keytab" Is Displayed.....	348
21.4 Worker Logs Are Empty After the Storm Topology Is Submitted.....	349
21.5 Worker Runs Abnormally After the Storm Topology Is Submitted and Error "Failed to bind to XXX" Is Displayed.....	351
21.6 "well-known file is not secure" Is Displayed When the jstack Command Is Used to Check the Process Stack.....	353
21.7 Data Cannot Be Written to Bolts When the Storm-JDBC Plug-in Is Used to Develop Oracle Databases.....	355
21.8 Internal Server Error Is Displayed When the User Queries Information on the Storm UI.....	357
<b>22 Using Ranger.....</b>	<b>359</b>
22.1 After Ranger Authentication Is Enabled for Hive, Unauthorized Tables and Databases Can Be Viewed on the Hue Page.....	359
<b>23 Using Yarn.....</b>	<b>361</b>

23.1 A Large Number of Jobs Occupying Resources After Yarn Is Started in a Cluster.....	361
23.2 Error "GC overhead" Is Reported When Tasks Are Submitted Using the hadoop jar Command on the Client.....	362
23.3 Disk Space of a Node Is Used Up Due to Oversized Aggregated Logs of Yarn.....	363
23.4 Temporary Files Are Not Deleted When a MapReduce Job Is Abnormal.....	365
23.5 Incorrect Port Information of the Yarn Client Causes Error "connection refused" After a Task Is Submitted.....	366
23.6 "Could not access logs page!" Is Displayed When Job Logs Are Queried on the Yarn Web UI.....	367
23.7 Error "ERROR 500" Is Displayed When Queue Information Is Queried on the Yarn Web UI.....	368
23.8 Error "ERROR 500" Is Displayed When Job Logs Are Queried on the Yarn Web UI.....	369
23.9 An Error Is Reported When a Yarn Client Command Is Used to Query Historical Jobs.....	370
23.10 Number of Files in the TimelineServer Directory Reaches the Upper Limit.....	371
<b>24 Using ZooKeeper.....</b>	<b>373</b>
24.1 An Error Is Reported When the MRS Client Is Used to Connect to ZooKeeper.....	373
24.2 ZooKeeper Is Unavailable Because of Non-synchronized Time Between Active and Standby Master Nodes.....	374
<b>25 Storage-Compute Decoupling.....</b>	<b>376</b>
25.1 A User Without the Permission on the /tmp Directory Failed to Execute a Job for Accessing OBS...	376
25.2 When the Hadoop Client Is Used to Delete Data from OBS, It Does Not Have the Permission for the .Trash Directory.....	377
25.3 An MRS Cluster Fails Authentication When Accessing OBS Because the NTP Time of Cluster Nodes Is Not Synchronized.....	379

# 1 Account Passwords

---

## 1.1 Resetting or Changing the Password of Manager User admin

- If you know the password, change the password by following the instructions in [Changing the Password for User admin](#).
- If you forget the password, reset the password by following the instructions in [Changing the Password of a Component Running User](#).

## 1.2 Failed to Download Authentication Credentials If the Username Is Too Long

### Issue

In MRS clusters 3.0.2 to 3.1.0, a maximum of 32 characters are allowed in the username when a user is added. However, if the username contains more than 20 characters, the user fails to download the Keytab file, and status code "400 Bad Request" is displayed.

### Symptom

In MRS clusters 3.0.2 to 3.1.0, a maximum of 32 characters are allowed in the username when a user is added. However, if the username contains more than 20 characters, the user fails to download the Keytab file, and status code "400 Bad Request" is displayed.

### Cause Analysis

The **validate-common-config.xml**, **validate-rule-session.xml**, and **validate-rule-user.xml** configuration files in the **/opt/Bigdata/om-server\_\*/apache-tomcat-\*/webapps/web/WEB-INF/validate** directory of the master node are incorrect and need to be modified.

## Procedure

**Step 1** Log in to the master node as user **omm** and switch to the **/opt/Bigdata/om-server\_\*/apache-tomcat-\*/webapps/web/WEB-INF/validate** directory.

```
cd /opt/Bigdata/om-server_*/apache-tomcat-*/webapps/web/WEB-INF/validate
```

**Step 2** Modify the **validate-common-config.xml** file.

```
vi validate-common-config.xml
```

Change the **maxLength** value of the username from **32** to **64**.

```
<!-- Username -->
<validators alias="USER_NAME">
  <validator name="RANGE_LENGTH_VALIDATOR" minLength="3"
    maxLength="64" />
  <validator name="REGEXP_VALIDATOR" rule="^[a-zA-Z0-9\ ]+$"
</validators>
```

**Step 3** Modify the **validate-rule-session.xml** file.

```
vi validate-rule-session.xml
```

Change the **rule** value from **20** to **64**.

```
<!-- Download the credentials of the current user -->
<param_validator url="/api/v2/session/user/keytab/download" method="get"
errorHandler="com.huawei.bigdata.om.web.api.validate.SpecialValidatorErrorHandler" dataPattern="form">
  <!-- Parameter name: File name -->
  <!--Validation rule: userName_13-digit number_keytab.tar; case sensitive-->
  <parameter name="file_name" required="true" errorKey="13-4000005"
errorMessage="RESID_OM_API_SESSION_0013">
    <validator name="REGEXP_VALIDATOR" rule="[\\-\\w ]{3,64}_d{13}_keytab\\.tar"
caseSensitive="true" />
  </parameter>
```

**Step 4** Modify the **validate-rule-user.xml** file.

```
vi validate-rule-user.xml
```

Change the **rule** value from **20** to **64**.

```
<!--Download the user credentials -->
<param_validator url="/api/v2/permission/users/keytab/download" method="get"
errorHandler="com.huawei.bigdata.om.web.api.validate.SpecialValidatorErrorHandler" dataPattern="form">
  <!--Mandatory; userName_13-digit number_keytab.tar; case sensitive-->
  <parameter name="file_name" required="true" errorKey="12-4000005"
errorMessage="RESID_OM_API_AUTHORITY_0005">
    <validator name="REGEXP_VALIDATOR" rule="[\\-\\w ]{3,64}_d{13}_keytab\\.tar"
caseSensitive="true" />
  </parameter>
</param_validator>
```

**Step 5** Restart Tomcat and wait until the startup is successful.

1. Run the following command as user **omm** to query the PID of the Tomcat process:

```
ps -ef|grep apache-tomcat
```

2. Run the **kill -9 PID** command to forcibly stop the specified Tomcat process. For example:

```
kill -9 1203
```

3. Run the following command to restart Tomcat:

```
sh ${BIGDATA_HOME}/om-server/tomcat/bin/startup.sh
```

**Step 6** Download the authentication credentials again.

**----End**

# 2 Account Permissions

## 2.1 A Message Is Displayed Indicating That the User Does Not Have the Permission to Obtain the MRS Cluster Hosts

### Issue

When a user calls the `/v1.1/<project_id>/clusters/<cluster_id>/hosts` API using the AK/SK to obtain hosts in an MRS cluster, message "User do not have right to access cluster" is displayed.

### Cause Analysis

Parameters such as **project\_id** are not filled in the request header. As a result, the **project\_id** of the token parsed by the cloud service is inconsistent with that of the cluster.

### Procedure

Before calling interfaces in AK/SK-based authentication mode, collect required information by referring to [Table 2-1](#). For details about how to sign requests and use SDKs, see [API Request Signing Guide](#).

Before constructing an API request, obtain the following information, including the endpoint and URI of the request URL, AK/SK used for signature and authentication, and project ID used to distinguish tenants.

**Table 2-1** Information to be collected

<b>Endpoint</b>	Endpoint of a cloud service in a region.
<b>Project_Id</b>	Project ID, which needs to be configured in the URI of most APIs to identify different projects.



<b>AK/SK</b>	Access key pair, including the access key ID (AK) and secret access key (SK), used to sign API requests.
<b>URI</b>	API request path and parameters. For details, see <a href="#">API Overview</a> .
<b>X-Domain-Id</b>	Account ID, which is used for: <ul style="list-style-type: none"><li>• Obtaining a token for token-based authentication.</li><li>• Calling APIs of global services using AK/SK-based authentication. <b>X-Domain-Id</b> needs to be configured in the request header.</li></ul>
<b>X-Project-Id</b>	Sub-project ID, which is used in multi-project scenarios. The <b>X-Project-Id</b> field is mandatory in the request header for accessing resources in a sub-project through AK/SK-based authentication.

 NOTE

For details about how to obtain the parameters in [Table 2-1](#), see [AK/SK Signing and Authentication Guide](#).

## 2.2 Failed to View MRS Cluster Details

### Symptom

A user fails to access the cluster details page by clicking the cluster name on the MRS console.

### Cause Analysis

1. The user's MRS cluster uses enterprise project A (with **MRS FULLACCESS** and **ECS FULLACCESS** permissions).
2. The VPC uses enterprise project B.
3. The security group uses enterprise project A.
4. No permissions are configured for the IAM user group.
5. The analysis shows that enterprise project B used by the user's VPC does not have the **vpc readonly** permission. As a result, the error is reported.

### Procedure

Add the **vpc readonly** permission to the enterprise project used by the VPC.

# 3 Common Exceptions in Logging In to the Cluster Manager

---

## 3.1 Failed to Access Manager of an MRS Cluster

### Symptom

After an MRS cluster is created, its Manager cannot be accessed from the console.

### Fault Locating

1. Check the MRS cluster status. If the cluster is stopped or being created, or an active/standby switchover is being performed, the Manager page cannot be accessed.
2. Check the access permission of the current user. **ReadOnlyAccess** indicates the read-only permission. Users with this permission can only view MRS resources but cannot access the Manager page.
3. Check whether an available EIP is bound. Manager can be accessed from a local PC only after an available EIP is bound.
4. Check the local user IP address in the security group rules. In the security group rules of the MRS cluster, port 9022 needs to be opened for the current user.
5. Check the local browser. For example, check whether the internal network proxy is configured and whether security settings that block user tokens are added.
6. Check the Manager running status. If network configurations such as the EIP and security group are correct, check whether Manager is running properly.
7. Check the user password. The user password used for logging in to Manager must be correct or valid.

### Procedure

- Step 1** Log in to the MRS management console, choose **Active Clusters**, and check whether the MRS cluster is running properly.

- If the MRS cluster is running properly, go to [Step 3](#).
- Otherwise, find the cause of this problem. If the MRS cluster is abnormal, rectify the cluster fault and access Manager again. If the MRS cluster is being created or an active/standby switchover is being performed, wait until the cluster creation or active/standby switchover is complete and access Manager again.

**Step 2** Check whether the permission configurations of the current user are correct and whether the user has the permission to access Manager.

For details, see [Creating an MRS User](#).

Do not add the MRS **ReadOnlyAccess** permission, which allows users only to view MRS resources but not to access Manager.

**Step 3** Click the desired cluster in the cluster list to go to the cluster details page.

**Step 4** Check the bound EIP.

Click **Access Manager** next to **MRS Manager** and select an available EIP from the EIP drop-down list.

 **NOTE**

- If no EIP is available, click **Manage EIP** to create one and select an EIP from the **EIP** drop-down list.
- If an EIP has been created but cannot be found during binding, the EIP may have been bound to another cluster. In this case, unbind the EIP on the **EIPs** page and then bind it to the current cluster.

**Step 5** Check whether the security group configurations are correct.

If Manager still cannot be accessed after the EIP is bound, check whether port 9022 is enabled in the security group rule of the current cluster.

1. On **Security Groups**, select the security group to which the security group rule to be added belongs.

 **NOTE**

You can view the security group name in **Security Group** on the **Dashboard** page of the cluster.

2. Click **Manage Security Group Rule** in **Add Security Group Rule** to add a security group rule. On the displayed page, click **Inbound Rules**. Check whether there is a security group rule in which the source address is the IP address of the current user to access the cluster, the protocol & port is **TCP : 9022**, and the action is **Allow** in the rule list. If not, click **Add Rule** to add such a rule and enable port 9022.

For example, if the current user accesses the cluster via IP address **10.x.x.x**, the security group rule is as follows.

**Figure 3-1** Adding a security group



 **NOTE**

The IP address that needs to be allowed in the security group is the public IP address used by users to access the MRS cluster. If the current network is an internal LAN, check that the IP address configured in the security group is the public network egress IP address.

3. Click **Add Security Group Rule** on the MRS **Dashboard** page to check whether the configured IP address is changed. If yes, change the IP address and access Manager again.

**Step 6** If the EIP and security group are normal, log in to Manager again and check whether the login page is displayed.

Enter the username and password and click **Login**.

 **NOTE**

- To log in as user **admin**, enter the password set during the cluster creation.
- To log in using the username and password created by the administrator, reset the password after the first login.
- If an error message is displayed indicating that the authentication information is invalid during the login, the password may be incorrect or has expired. Enter the correct password, or change the password by referring to [Resetting or Changing the Password of Manager User admin](#).
- Check whether the internal network proxy is configured and whether security settings that can intercept user tokens are added in the browser.
- If the Manager login page cannot be displayed after multiple updates, clear the browser cache or use another browser.

If the Manager page is still blank or abnormal after a successful login, rectify the fault by referring to [Accessing the Web Pages](#).

**Step 7** After logging in to Manager, access the web UIs of MRS cluster components.

On Manager, choose **Cluster > Services** and click the desired component name. On the overview page, log in to the web UIs of MRS components.

For example, choose **Cluster > Services** and click **HDFS**. On the overview page, click the hyperlink next to **NameNode WebUI** to access the HDFS web UI.

If a blank page or an exception occurs when the web UI of a component is accessed, rectify the fault by referring to [Common Exceptions in Accessing the MRS Web UI](#).

----End

## 3.2 Accessing the Web Pages

### 3.2.1 Error "502 Bad Gateway" Is Reported During the Access to MRS Manager

#### Issue

Error **502 Bad Gateway** is reported when a user attempts to access MRS Manager.

## Symptom

In an MRS cluster with Kerberos authentication disabled, an error message is displayed when you access the MRS Manager page.

Figure 3-2 Error message



## Cause Analysis

The MRS domain name is changed from console-emr to mrs. As a result, the link between the common cluster and MRS Manager is incorrect.

## Procedure

**Step 1** Log in to each Master node as user **root**.

**Step 2** Go to the **/opt/knox/conf/** directory and find the **ext.properties** file.

```
[root@node-~]# cd /opt/knox/conf/
[root@node-~]# ls
descriptors          gateway-site.xml      krb5JAASLogin.conf   shared-providers     user.keytab
ext.properties       knoxcli-log4j.properties  ldap-log4j.properties  shell-log4j.properties  users.ldif
gateway-log4j.properties  krb5.conf             README                topologies
```

**Step 3** Change the value of **console-emr** in the **ext.properties** file on all Master nodes to **mrs**.

```
# Split by comma, exclude those cookies in request
exclude_request_cookies=newCluster,preCluster

# Console context path
console_context=console-emr
```

Change to "mrs" for all master nodes.

**Step 4** Go to the **/opt/knox/bin/** directory and run the **su - omm** command to switch to user **omm**.

**Step 5** Run the **restart-knox.sh** script to restart the Knox service.

**Step 6** Try to access MRS Manager again.

----End

## 3.2.2 An Error Message Occurs Indicating that the VPC Request Is Incorrect During the Access

### Symptom

An error message indicating that the VPC request is incorrect is displayed when Manager is accessed on the console.

## Possible Causes

- The current user lacks the permission to access the VPC.
- The group to which the IAM sub-user belongs has mutually exclusive permissions. As a result, Manager cannot be accessed.
- The current user lacks the permission to add IP addresses to the security group, or the IP addresses added to the security group have changed.

## Procedure

### Step 1 Set user permissions.

For how to create an MRS user, see [Creating an MRS User](#). Alternatively, add user permissions. Do not add the MRS **ReadOnlyAccess** permission. Users have this permission can only view MRS resources.

#### NOTE

By default, new IAM users do not have any permissions. To assign permissions to a user, add the user to one or more groups and assign permissions policies or roles to these groups. The user then inherits permissions from the groups it is a member of and can perform specified operations on cloud services based on the permissions.

### Step 2 On the MRS console, choose **Clusters > Active Clusters**. Click the desired cluster in the cluster list to access its details page. Check that the local IP address is correctly configured in the security group.

- MRS 3.x or later versions: For details, see the "Accessing FusionInsight Manager Using EIP" part of "Accessing Manager" > "Accessing FusionInsight Manager (MRS 3.x or later )" in *User Guide*.
- Versions before MRS 3.x: For details, see the "Accessing FusionInsight Manager Using an EIP" part of "Accessing Manager" > "Accessing MRS Manager (MRS 2.x or Earlier )" in *User Guide*.

----End

## 3.2.3 Error 503 Is Reported When Manager Is Accessed Through Direct Connect

### Symptom

When Manager is accessed through Direct Connect, error 503 is displayed, indicating that the page cannot be used currently.

### Possible Causes

As the number of tasks increases or the number of concurrent tasks increases, the Executor memory may be insufficient. As a result, the Manager cannot be accessed.

### 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 commands to modify the **catalina.sh** script:  
**vim /opt/executor/bin/catalina.sh**  
Search for **JAVA\_OPTS**, find the configuration similar to **JAVA\_OPTS="-Xms1024m -Xmx4096m"**, change the value as required, and save the modification.
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.shsh /opt/executor/bin/startup.sh**
4. Log in to the Manager page again.

### 3.2.4 Error Message "You have no right to access this page." Is Displayed When Users log in to the Cluster Page

#### Symptom

The following error message is displayed when you access the cluster management page.

You have no right to access the page.

#### Fault Locating

- The currently logged-in user does not have access to FusionInsight Manager.
- The disk space of the partition where **/srv/BigData/dbdata\_om** is located is used up.

#### Procedure

**Step 1** Log in to the FusionInsight Manager page as user **admin** to check the permissions of the current user.

**Step 2** The user must have at least one of the following permissions for accessing the FusionInsight Manager page.

<input type="checkbox"/>	Manager_administrator	▶ Manager system administrator: has all rights of Manager system ...
<input checked="" type="checkbox"/>	Manager_auditor	▶ Manager system auditor: has rights to view and manager audit inf...
<input type="checkbox"/>	Manager_operator	▶ Manager system operator: has all rights of Manager system exce...
<input type="checkbox"/>	Manager_tenant	▶ Manager tenant viewer: has rights to view tenant.
<input type="checkbox"/>	Manager_viewer	▶ Manager system viewer: has rights to view dashboard, services, ...

**Step 3** Log in to the active OMS node as user **rootoromm**.

**Step 4** Check whether the partition where **/srv/BigData/dbdata\_om** resides is used up.

**Step 5** Clear unnecessary data.

----End

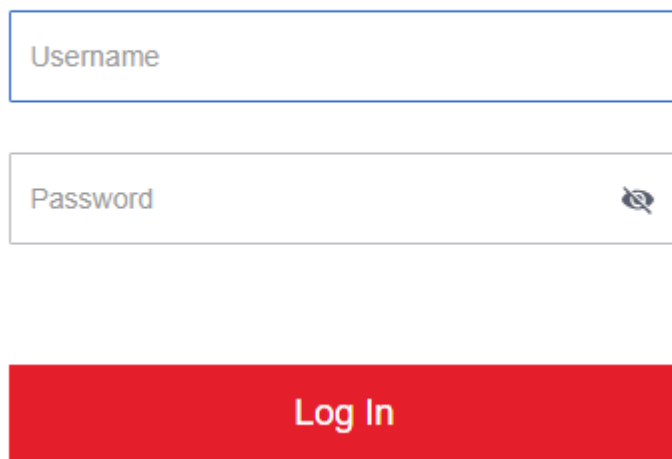
## 3.2.5 Error Message "Invalid credentials" Is Displayed When a User Logs In to Manager

### Symptom

Error message "Invalid credentials" is displayed when a user logs in to Manager.

### Login

Invalid credentials.



The screenshot shows a login interface with the following elements:

- A header "Login" in bold black text.
- An error message "Invalid credentials." in red text.
- A text input field labeled "Username" with a blue border.
- A text input field labeled "Password" with a blue border and a small eye icon on the right side.
- A large red button labeled "Log In" in white text.

### Possible Causes

- The username or password is not correct.
- The password of the current login user has expired.

### Procedure

**Step 1** If the password is incorrect, check and enter the correct password.

(The default administrator of the MRS cluster is **admin** and the password is set when the cluster is created.)

Alternatively, reset the user password by referring to [Resetting or Changing the Password of Manager User admin](#).

**Step 2** If the password of a cluster user expires or you forgot the password, log in to Manager as an administrator to initialize the password.

1. Log in to FusionInsight Manager.
2. Choose **System** > **Permission** > **User**.



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**, and enter the password.

After the password is reset, you need to change the password again when using the new password to log in to Manager.

----End

## 3.2.6 Failed to Log In to the Manager After Timeout

### Symptom

Login occasionally fails after MRS Manager exits due to timeout.

### Possible Causes

After the timeout, the JS and CSS files of the page fail to be loaded.

### Procedure

**Step 1** Clear the browser cache.

**Step 2** Refresh the current page and log in again.

----End

## 3.2.7 Failed to Log In to MRS Manager After the Python Upgrade

### Issue

Failed to log in to MRS Manager after Python is upgraded.

### Symptom

After Python is upgraded, MRS Manager fails to be accessed using the **admin** account and the correct password.

### Cause Analysis

When upgrading Python to Python 3.x, the user modifies the file directory permission of **openssl**. As a result, the LdapServer service cannot be started, causing a login authentication failure.

### Procedure

**Step 1** Log in to the Master node in the cluster as user **root**.

**Step 2** Run the **chmod 755 /usr/bin/openssl** command to modify the file directory permission of **/usr/bin/openssl** to **755**.

**Step 3** Run the **su omm** command to switch to user **omm**.

**Step 4** Run the **openssl** command to check whether the **openssl** mode can be entered.

If it can be entered, the permission has been modified successfully. If it cannot be entered, the permission fails to be modified.

If the permission fails to be modified, check whether the command is correct or contact O&M personnel.

**Step 5** After the permission is modified, the LdapServer service will be restarted. After the LdapServer service is restarted, log in to MRS Manager again.

----End

## 3.2.8 Failed to Log In to MRS Manager After Changing the Domain Name

### Symptom

After changing the domain name, the user cannot log in to MRS Manager through the console, or fails to log in to MRS Manager.

### Cause Analysis

After the domain name is changed, the **keytab** file of user **executor** is not updated. As a result, the executor process repeatedly performs authentication after the authentication fails, causing memory overflow of the ACS process.

### Procedure

**Step 1** Restart the ACS process.

1. Log in to the active management node (master node marked a solid star on the **Nodes** tab of the MRS cluster) as user **root**.
2. Run the following commands to restart the acs process:  
**su - omm**  
**ps -ef|grep =acs** (Query the PID of the acs process.)  
**kill -9 PID** (Replace *PID* with the acs process ID to kill the acs process.)
3. Wait for several minutes and run the **ps -ef|grep =acs** command to check whether the acs process is automatically started.

**Step 2** Replace the **keytab** file of user **executor**.

1. Log in to MRS Manager and choose **System > User**. In the **Operation** column where user **executor** resides, click **Download Authentication Credential**. Decompress the package to obtain the **keytab** file.
2. Log in to the active management node as user **root** and replace the **/opt/executor/webapps/executor/WEB-INF/classes/user.keytab** file with the file obtained in [Step 2.1](#).

**Step 3** Replace the **keytab** and **conf** files of user **knox**.

1. Log in to MRS Manager and choose **System > User**. In the **Operation** column where user **knox** resides, click **Download Authentication Credential**. Decompress the package to obtain the **keytab** and **conf** files.
2. Log in to the active management node as user **root** and replace the **/opt/knox/conf/user.keytab** with the file obtained in [Step 3.1](#).
3. Change the **principal** value in the **/opt/knox/conf/krb5JAASLogin.conf** file to the new domain name.
4. Replace the **/opt/knox/conf/krb5.conf** file with the **krb5.conf** file obtained in [Step 3.1](#).

**Step 4** Back up the original client directory.

```
mv {Client directory} /opt/client_init
```

**Step 5** Reinstall the client. For details, see [Updating a Client](#).

**Step 6** Log in to the active and standby management nodes as user **root** and run the following commands to restart the knox process:

```
su - omm
```

```
ps -ef | grep gateway | grep -v grep (Search for the PID of the knox process.)
```

```
kill -9 PID (Replace PID with the ID of the knox process to kill the knox process.)
```

```
/opt/knox/bin/restart-knox.sh (Start the knox process.)
```

**Step 7** Log in to the active and standby management nodes as user **root** and run the following commands to restart the executor process:

```
su - omm
```

```
netstat -anp |grep 8181 |grep LISTEN (Search for the PID of the executor process.)
```

```
kill -9 PID (Replace PID with the ID of the executor process to kill the executor process.)
```

```
/opt/executor/bin/startup.sh (Start the executor process.)
```

```
----End
```

## 3.2.9 Manager Page Is Blank After a Success Login

### Issue

After a user logs in to FusionInsight Manager, the page displayed is blank.

### Symptom

After a user logs in to FusionInsight Manager, the page displayed is blank.

### Possible Causes

- Login to the Manager page fails, and the browser cache needs to be cleared.
- The permission on the **/opt/Bigdata/tmp** directory on the Master node is modified.

## Procedure

### Clearing the browser cache

**Step 1** Open the browser (using Google Chrome as an example), and press **Ctrl+Shift+Delete**. The dialog box for clearing browsing data is displayed.

**Step 2** Select the browsing records to be cleared and click **Clear Data**.

**Step 3** Check whether the Manager page is accessible.

- If yes, no further action is required.
- If no, go to [Step 4](#).

Modifying the permission on the **/opt/patch\_install** directory

**Step 4** Log in to the Master1 and Master2 nodes as user **root**.

**Step 5** Check whether the permission on the **/opt/Bigdata/tmp** directory of the Master1 and Master2 nodes is modified. (The default permission is 770.)

**Step 6** If the permission is modified, run the following command to restore the permission on the **/opt/Bigdata/tmp** directory:

```
chmod 1770 /opt/Bigdata/tmp
```

**Step 7** Log in to the active OMS node and run the following command to restart the node:

```
sh ${BIGDATA_HOME}/om-server/om/sbin/restart-oms.sh
```

The command is run successfully if the following information is displayed:

```
start HA successfully.
```

**Step 8** Wait for a while and log in to Manager again.

----End

## 3.2.10 Cluster Login Fails Because Native Kerberos Is Installed on Cluster Nodes

### Symptom

Users fail to log in to the cluster or change the password after the MRS cluster has been installed.

### Cause Analysis

1. Log in to the Kerberos service node of the cluster and view **/var/log/Bigdata/kerberos/krb5kdc.log**.
2. Log in to the active and standby OMS nodes and view **/var/log/Bigdata/okerberos/oms-krb5kdc.log**.

It is found that there are lots of authentication failures and error message "<unknown client> for <unknown server>" is displayed.

3. According to the check, the authentication exception is caused by the native Kerberos installed on the node.

Run the `rpm -qa | grep krb5` command to check whether `krb5-workstation-1.10.3-33.el6.x86_64` exists.

## Procedure

**Step 1** Run the following command to uninstall the software and its dependency software.

```
rpm -e krb5-workstation-1.10.3-33.el6.x86_64
```

**Step 2** On the node where Kerberos is installed, run the `kill -9` command to stop the Kerberos process. Wait until the Kerberos process is restarted.

----End

## 3.2.11 Using Google Chrome to Access MRS Manager on macOS

### Issue

The user failed to access MRS Manager using Google Chrome on macOS.

### Symptom

MRS Manager is inaccessible from Google Chrome on macOS.

### Possible Causes



Access failed due to restrictions of the Mac certificates.

## Procedure

**Step 1** Enter the login URL of MRS Manager in the address box of Google Chrome.

The MRS Manager login page is displayed, although Google Chrome cannot be used to log in to MRS Manager at this time. Proceed with the next step.

**Step 2** Search for the certificate.

Click  in the upper left corner of the browser and click  in the target row.

**Step 3** Obtain the certificate.

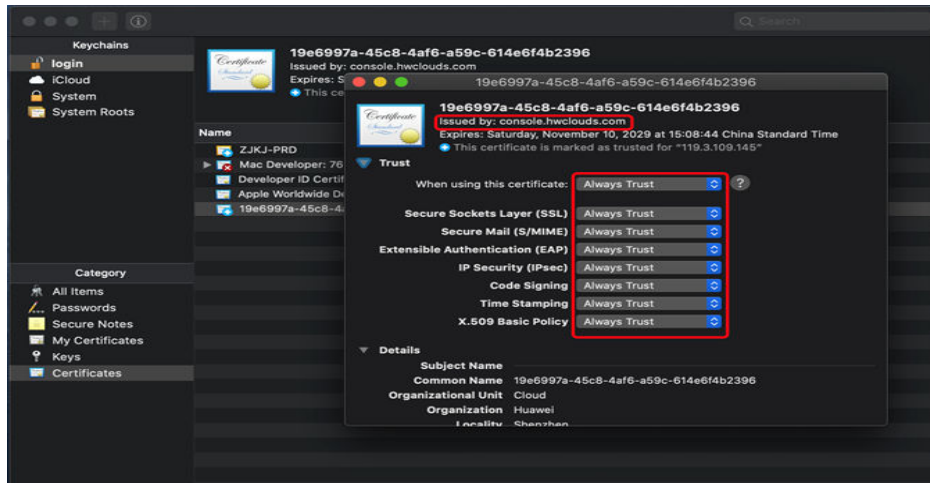
Select the certificate and drag it to the desktop to obtain the certificate.

**Step 4** In the Mac application, choose **Keychain Access**.

**Step 5** Set **Keychains** to **login** and **Category** to **Certificates**.

**Step 6** Drag the certificate obtained in **Step 3** to the keychain list.

**Step 7** Double-click the certificate, configure **Always Trust** for it, and close the window.



**Step 8** In the displayed dialog box, enter the login password. Then you can access MRS Manager using Google Chrome.

----End

## 3.2.12 How Do I Unlock a User Who Logs in to Manager?

### Symptom

1. When logging in to Manager, the system displays a message indicating that the account is locked. Contact the system administrator to obtain the access permission again.
2. The user who logs in to the Manager page is locked. As a result, jobs cannot be submitted and the Manager page cannot be accessed.
3. The following error message is displayed when the **kinit user** command is executed when a job is submitted in the background:  
kinit: Clients credentials have been revoked while getting initial credentials

### Cause Analysis

The number of incorrect password attempts for logging in to Manager exceeds the upper limit, and the account is locked.

### Procedure

**Step 1** Log in to the **kadmin** console.

1. Log in to the node where the client is installed as the **root** user and run the following command to switch to the client directory, for example, **/opt/client**.  
**cd /opt/client**
2. Run the following command to configure environment variables.  
**source bigdata\_env**
3. Run the following command and enter the password of user **kadmin/admin** to log in to the **kadmin** console:  
**kadmin -p kadmin/admin**

 NOTE

In 3.x and later versions, the default password is **Admin@123**. In versions earlier than 3.x, the default password is **KAdmin@123**.

After the first login, the system displays a message indicating that the password has expired. Change the password as prompted and keep it secure.

- Step 2** Run the following command to check whether the value of **Failed password attempts** is greater than or equal to 5. If yes, the account is locked.

```
getprinc Locked User Name
```

- Step 3** If the account is locked, run the following command to unlock the account and log in to Manager again:

```
modprinc -unlock Locked User Name
```

```
----End
```

### 3.2.13 Why Does the Manager Page Freeze?

#### Question

The Manager page of MRS clusters responds slowly.

#### Symptom

After a user logs in to Manager of an MRS cluster and clicks a button, data can be loaded only one minute later. After some buttons are clicked, frame freezing occurs and the pages cannot be accessed.

#### Possible Causes

View the `/var/log/Bigdata/omm/oms/pms/scriptlog/pms_script.log` file. The content of the `/var/log/Bigdata/omm/oms/pms/pms.log` file is similar to the following:

```
# java.lang.OutOfMemoryError: GC overhead limit exceeded
# -XX:OnOutOfMemoryError="echo [ERR00]: JVM OutMemory >> /var/log/Bigdata/omm/oms/pms/scriptlog/pms_script.log;kill -9 zp"
# Executing /bin/sh -c "echo [ERR00]: JVM OutMemory >> /var/log/Bigdata/omm/oms/pms/scriptlog/pms_script.log ..."
[ERR00]: JVM OutMemory
# Executing /bin/sh -c "kill -9 25000"..."
/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/bin/omm_s_pm_ctl.sh: line 334: 25000 Killed                  java -Dprocess.nu
sd -Xms1024m -Xmx1024m -Xss256k -XX:OnOutOfMemoryError="echo [ERR00]: JVM OutMemory >> /var/log/Bigdata/omm/oms/pms/script
s_script.log;kill -9 zp" -XX:ErrorFile=/var/log/Bigdata/omm/oms/pms/pmsdump/pms_error_211129223734.log -Djava.net.preferIP
k=true -Djava.net.preferIPv6Addresses=false -Dbase.dir=/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib -cp /opt/Bigdata/
om-server-8.1.0.1/OMS/workspace/lib/snakeyaml-1.27.jar:/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib/c3p0-0.9.5.4.jar:/opt/Big
om-server-8.1.0.1/OMS/workspace/lib/quartz-2.3.2.jar:/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib/commons-collections4
ar:/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib/commons-io-2.8.0.jar:/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib/y
2B.0-je.jar:/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib/manager-pms-8.1.0-310013-All.jar:/opt/Bigdata/om-server-8.1.0
S/workspace/lib/manager-pms-8.1.0-310013-EXTEN.jar:/opt/Bigdata/om-server-8.1.0.1/OMS/workspace/lib/kafka-clients-2.4.0-ja
```

According to the logs, memory overflow occurs. Therefore, it can be determined that the Manager page freezes due to insufficient PMS memory.

#### Procedure

- Step 1** Log in to the active management node as user **omm**.
- Step 2** Run the following command to open the **application.properties** file:

```
vi ${BIGDATA_HOME}/om-server_*/OMS/workspace0/conf/pms/
application.properties
```

- Step 3** Increase the value of **pms.mem** for the PMS process based on site requirements, save the modification, and exit.

```
...  
pms.mem=800m
```

The increment of the new value cannot be greater than the free memory of the OS.

- Step 4** Run the following commands to query the PID of the PMS process, stop the PMS process, and wait for the process to restart:

```
ps -ef | grep pms
```

```
kill -9 PID
```

- Step 5** Log in to Manager and check whether frame freezing disappears.

----End

## 3.3 Common Exceptions in Accessing the MRS Web UI

### 3.3.1 How Do I Do If an Error Is Reported or Some Functions Are Unavailable When I Access the Web UIs of HDFS, Hue, YARN, HetuEngine, and Flink?

#### Question:

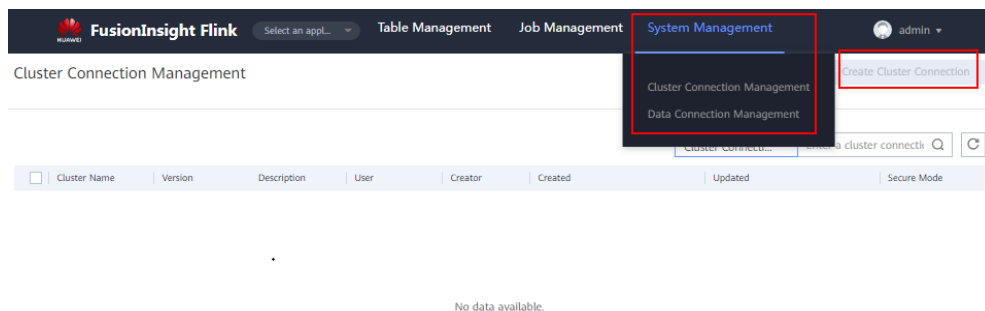
What can I do if an error is reported or some functions are unavailable when I access the WebUI of components such as HDFS, Hue, Yarn, Flink, and HetuEngine?

#### Answer:

Users who access the WebUI of components such as HDFS, Hue, Yarn, Flink, and HetuEngine do not have the management permission of the corresponding components. As a result, an error is reported or some functions are unavailable.

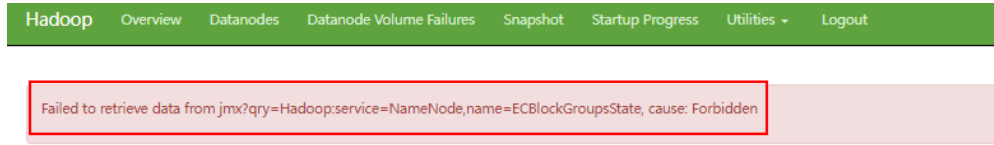
Example:

- After you log in to the web UI of Flink as the current user, some content cannot be displayed, and you do not have the permission to create applications, cluster connections, or data connections.



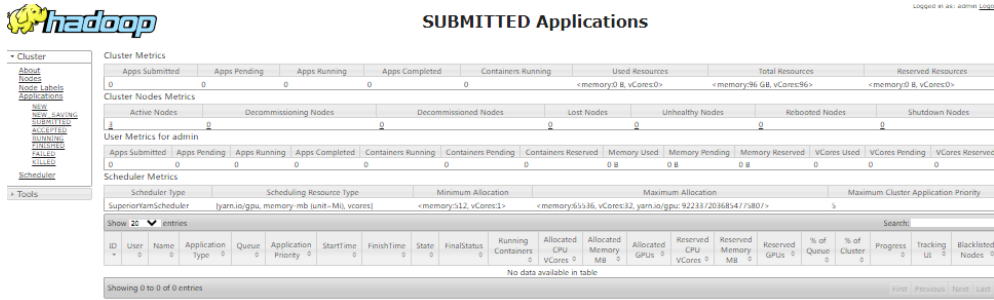



- After you log in to the web UI of HDFS as the current user, error message "Failed to retrieve data from /jmx?qry=java.lang:type=Memory, cause: Forbidden" is displayed.

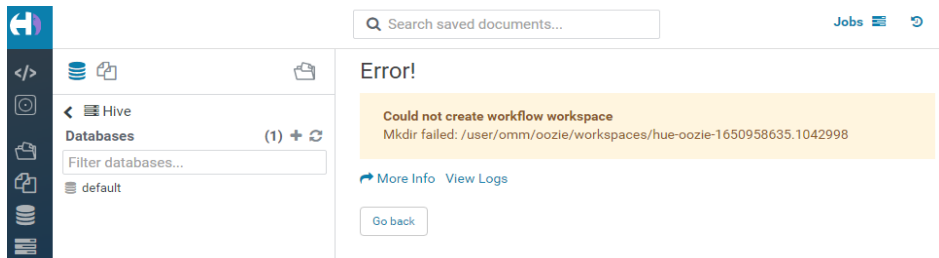


Hadoop, 2021.

- After you log in to the web UI of YARN as the current user, you cannot view job information.



- After you log in to the web UI of Hue as the current user, click  in the navigation pane on the left, and select **Workflow**, an error message is displayed.



You are advised to log in to the web UIs of the components as a user with corresponding management permissions. For example, you can create a service user who has the management permissions on HDFS and you can log in to the web UI of HDFS as the created user. For details, see [Creating a User](#).

### 3.3.2 Error 500 Is Reported When a User Accesses the Component Web UI

#### Issue

Error 500 is reported when a user accesses the component web UI.

#### Symptom

Error 500 is reported when a user accesses the component web UI. The detailed error information is as follows:

```
HTTP ERROR 500 java.io.IOException: Service connectivity error.
```

## Possible Causes

The possible cause is that the number of Knox threads is insufficient.

## 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** Wait for 5 minutes and check whether the component web UI can be accessed.

- If yes, no further action is required.
- If no, contact O&M personnel.

----End

## 3.3.3 [HBase WebUI] Users cannot switch from the HBase WebUI to the RegionServer WebUI

### Symptom

In an MRS 1.9.3 cluster, on the HBase WebUI, click any RegionServer name in the **ServerName** column on the **Base Status** tab page in the **Region Servers** area. The corresponding page cannot be displayed.

### Procedure

**Step 1** Log in to the node where the active OMS resides as user **root** and run the following command to obtain the URL of RegionServer:

```
cat /opt/Bigdata/Apache-httpd*/conf/whitelist.txt |grep RegionServer
```

```
[root@node-master1GGLT ~]# cat /opt/Bigdata/Apache-httpd*/conf/whitelist.txt |grep RegionServer
/HBase/RegionServer/44 https://node-ana-coreqacr0001.mrs-11ty.com:16030
/HBase/RegionServer/45 https://node-ana-coreqacr0002.mrs-11ty.com:16030
[root@node-master1GGLT ~]#
```

**Step 2** Change **/HMaster/ID/master-status** to **/RegionServer/ID/rs-status** to access the RegionServer information page. In the preceding command, **ID** in **/RegionServer/ID/rs-status** is the **ID** corresponding to the RegionServer host name queried in **Step 1**.

For example, if the URL of the RegionServer is `/HBase/RegionServer/44`, the access address is `https://100.94.xxx.xxx:9022/component/HBase/RegionServer/44/rs-status`.

----End

### 3.3.4 [HDFS WebUI] When users access the HDFS WebUI, an error message is displayed indicating that the number of redirections is too large

#### Symptom

Switching to the HDFS WebUI from the Manager page occasionally fails, and an error message is displayed indicating that the number of redirections is too large. In this case, the WebUIs of other components can be accessed normally.

#### Cause Analysis

Knox is an active-active component. During the running of Knox, data is not synchronized between the active and standby nodes.

#### Procedure

- Step 1** Redirect to the corresponding native page through any other component page and wait until the redirection is successful.
- Step 2** Access the HDFS native page.

----End

### 3.3.5 [HDFS WebUI] Failed to access the HDFS WebUI using the Internet Explorer

#### Issue

Occasionally, Internet Explorer 9, Explorer 10, or Explorer 11 fails to access the native HDFS UI.

#### Symptom

The page fails to be accessed, and the browser cannot display the page, as shown in the following figure.



Turn on TLS 1.0, TLS 1.1, and TLS 1.2 in Advanced settings and try connecting to

#### Cause Analysis

Some earlier versions of Internet Explorer fail to process SSL handshake.

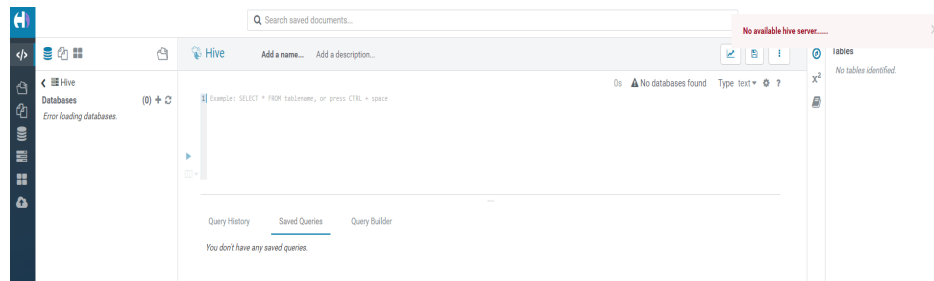
## Procedure

Refresh the page or use Google Chrome to access the page.

### 3.3.6 [Hue Web UI] A "No Permission" Error Is Displayed When a User Log In to the Hue Web UI

#### Symptom

In a cluster with Kerberos authentication enabled, when a user logs in to the Hue web UI as user **admin**, an error message is displayed, indicating that the user does not have the permission or cannot connect to the Hive server. The system may remind you to check whether the Hive permission of the current user and the Hive configuration of Hue are correct or display message "No available hive server...".



#### Possible Causes

For clusters with Kerberos authentication enabled, user **admin** does not have the permission to perform operations the Hue web UI. To solve this problem, you can create a human-machine user (for example, **hueuser**) and add it to the **hive**, **hadoop**, and **supergroup** user groups and the **System\_administrator** role. The primary group is **hive**.

## Procedure

- Step 1** Log in to FusionInsight Manager.
- Step 2** Choose **System > Permission > User**.
- Step 3** On the **User** page, click **Create**, configure the following information, and retain the default values for other parameters.
  - **Username:** user-defined, for example, **hueuser**.
  - **User Type:** Select **Human-Machine**.
  - **Password** and **Confirm Password:** Enter a password as prompted.
  - **User Group:** Click **Add**, select **hive**, **hadoop**, and **supergroup**, and click **OK**.
  - **Primary Group:** Select **hive**.
  - **Role:** Click **Add** and select **System\_administrator**.
- Step 4** Click **OK**.
- Step 5** Access the Hue web UI as the new user (for example, user **hueuser**).

Change the password upon the first login.

----End

### 3.3.7 [Hue Web UI] Failed to Access the Hue Web UI

#### Symptom

The Hue web UI page fails to be opened.

#### Possible Causes

- The port is not opened for external users.
- The floating IP address is not correctly set.
- Concurrent mass data query prolongs the query process, occupying the Hive link in Hue. As a result, Hue access is abnormal.

#### Fault Handling

1. On the local server where the browser locates, run the **telnet** command in the command-line interface (CLI) to check whether the port is opened.
2. Run **ifconfig** in the node where the Hue locates to check whether the floating IP address that is set takes effect.
3. It takes about 10 minutes to rectify the fault.
4. Check whether a large amount of data is queried at a time.

#### Procedure

**Step 1** Log in to the FusionInsight Manager portal.

**Step 2** Click **Cluster**, click the name of the desired cluster, and choose **Service > Hue**. On the displayed page, click **Instance** to check whether the active/standby status of the Hue service is normal.

- If the active/standby status is normal, go to [Step 7](#).
- If the active/standby status is abnormal, go to [Step 3](#).

**Step 3** Click **Configurations** and check whether configuration items **HUE\_FLOAT\_IP** and **HTTP\_PORT** are correct.

- If the configuration items are correct, go to [Step 4](#).
- If the configuration items are incorrect, correct the floating IP address and port number of the Hue service.

**Step 4** Run the **telnet Hue node IP address HTTP\_PORT** or **telnet HUE\_FLOAT\_IP HTTP\_PORT** command on the service network to check whether the port is open to external systems.

- If the port is open to external systems, check whether the port is correct and log in to the web UI again.
- If the port is not open to external systems, go to [Step 5](#).

**Step 5** Choose **Cluster**, click the name of the desired cluster, and choose **Service > Hue**. On the displayed page, click **Stop Service**.

- Step 6** Log in to the management node as user **root** and run the **ping** *HUE\_FLOAT\_IP* command to check whether the IP address can be pinged.
- If the IP address can be pinged, the floating IP is in use. Configure another IP address.
  - If the IP address cannot be pinged, go to [Step 7](#).

- Step 7** Check whether a large amount of data is queried at a time.
- If yes, optimize Hue query to avoid concurrent mass data query.
  - If no, go to [Step 8](#).

- Step 8** Contact O&M personnel.

----End

### 3.3.8 [Hue WebUI] The error "Proxy Error" is reported when a user accesses the Hue WebUI

#### Symptom

"Proxy Error" is displayed when the Hue page is opened.

#### Proxy Error

The proxy server received an invalid response from an upstream server.  
The proxy server could not handle the request *GET /*.

Reason: **Error reading from remote server**

#### Cause Analysis

- Httpd does not receive packets or receives incorrect packets due to network faults, such as the high packet error rate and high latency.
- The authentication relationship between the Httpd service and the Hue service fails because of clock change or other reasons.

#### Fault Locating

1. Check whether a network fault or clock change occurs.
2. Check whether alarms related to the Hue service have been generated or still exist.

#### Procedure

- Step 1** Log in to FusionInsight Manager to check whether any network fault, clock change, or Hue service alarm exists. If yes, contact the administrator to rectify the fault. Then, check whether the Hue page can be normally opened.
- If yes, no further action is required.
  - If no, go to [Step 2](#).

- Step 2** Obtain the address of the current Httpd server, which is the host address in the URL of the Hue page. That is, if the URL of Hue is https://192.168.0.1:21201, the Httpd service address is 192.168.0.1.
- Step 3** Log in to the Httpd server and check logs in `/var/log/Bigdata/httpd/error_log`.
- Step 4** Check whether the error message "AH01102: error reading status line from remote server" or "AH00898: Error reading from remote server" exists in the `error_log` file. (You can open the Hue page for several times to confirm the error.)
- If yes, restart the Hue service on FusionInsight Manager and go to **Step 5**.
  - If no, go to **Step 6**.
- Step 5** Restart the Hue service and open the Hue page again.
- If yes, no further action is required.
  - If no, go to **Step 6**.
- Step 6** Contact the O&M personnel.
- End

### 3.3.9 [Hue WebUI] Why Is the Hue Native Page Cannot Be Properly Displayed If the Hive Service Is Not Installed in a Cluster?

#### Question

Why is the native Hue page blank if the Hive service is not installed in a cluster?

#### Answer

In MRS 3.x, Hue depends on Hive.

### 3.3.10 Hue (Active) Cannot Open Web Pages

#### Symptom

The following information is displayed on the web UI of Hue (active):

**Service Unavailable**

The server is temporarily unable to service your request due to maintenance downtime or capacity problems. Please try again later.

#### Cause Analysis

- The Hue configuration has expired.
- The configuration of the Hue service needs to be modified manually in a single-master cluster of MRS 2.0.1 or later.

#### Solution

- If the Hue configuration has expired, restart the Hue service.
- Manually modify the Hue service configuration for a single-master cluster of MRS 2.0.1 or later.

- a. Log in to the Master node.
- b. Run the **hostname -i** command to obtain the IP address of the local host.
- c. Run the following command to obtain the value of **HUE\_FLOAT\_IP**:  

```
grep "HUE_FLOAT_IP" ${BIGDATA_HOME}/MRS_Current/1_*/etc*/ENV_VARS,
```

Replace *MRS* with the actual file name.
- d. Check whether the local IP address is the same as the value of **HUE\_FLOAT\_IP**. If they are different, change the value of **HUE\_FLOAT\_IP** to the local IP address.
- e. Restart the Hue service.

### 3.3.11 [Ranger WebUI] Why Cannot a New User Log In to Ranger After Changing the Password?

#### Question

When a new user logs in to Ranger, why is the 401 error reported after the password is changed?

#### Answer

The UserSync synchronizes user data at an interval of 5 minutes by default. Therefore, a new user created on Manager cannot log in to the Ranger before the user data is successfully synchronized because the Ranger database does not have the user information. The user can log in to the Ranger only after the specified interval ends.

If Kerberos authentication is disabled, only the **admin** user can log in to the Ranger page because Ranger does not synchronize user data from Manager.

### 3.3.12 [Tez WebUI] Error 404 is reported when users access the Tez WebUI

#### Question

When a user logs in to Manager and switches to the Tez web UI, error 404 or 503 is displayed.

#### HTTP ERROR 404

Problem accessing /null/applicationhistory. Reason:

Not Found

Powered by Jetty:// 9.3.20.v20170531

Adapter operation failed Å» 503: Error accessing https://:20026/Yam/TimelineServer/57/ws/v1/timeline/TEZ\_DAG\_ID



## Answer

The Tez web UI depends on the TimelineServer instance of Yarn. Therefore, TimelineServer must be installed in advance and in the **Good** state.

### 3.3.13 [Spark WebUI] Why Cannot I Switch from the Yarn Web UI to the Spark Web UI?

#### Question

The Spark application is run in yarn-client mode on the client. The following error occurs during the switch from the Yarn web UI to the application web UI:

#### Error Occurred.

Problem accessing /proxy/application\_ [redacted] /

Powered by Jetty://

The YARN ResourceManager log shows the following information:

```
2016-07-21 16:35:27,099 | INFO | Socket Reader #1 for port xxx | Auth successful for mapred/hadoop.<System domain name>@<System domain name> (auth:KERBEROS) | Server.java:1388
2016-07-21 16:35:27,105 | INFO | 1526016381@qtp-1178290888-1015 | admin is accessing unchecked http://10.120.169.53:23011 which is the app master GUI of application_1468986660719_0045 owned by spark | WebAppProxyServlet.java:393
2016-07-21 16:36:02,843 | INFO | Socket Reader #1 for port xxx | Auth successful for hive/hadoop.<System domain name>@<System domain name> (auth:KERBEROS) | Server.java:1388
2016-07-21 16:36:02,851 | INFO | Socket Reader #1 for port xxx | Auth successful for hive/hadoop.<System domain name>@<System domain name> (auth:KERBEROS) | Server.java:1388
2016-07-21 16:36:12,163 | WARN | 1526016381@qtp-1178290888-1015 | /proxy/application_1468986660719_0045/: java.net.ConnectException: Connection timed out | Slf4jLog.java:76
2016-07-21 16:37:03,918 | INFO | Socket Reader #1 for port xxx | Auth successful for hive/hadoop.<System domain name>@<System domain name> (auth:KERBEROS) | Server.java:1388
2016-07-21 16:37:03,926 | INFO | Socket Reader #1 for port xxx | Auth successful for hive/hadoop.<System domain name>@<System domain name> (auth:KERBEROS) | Server.java:1388
2016-07-21 16:37:11,956 | INFO | AsyncDispatcher event handler | Updating application attempt appattempt_1468986660719_0045_000001 with final state: FINISHING, and exit status: -1000 | RMAAppAttemptImpl.java:1253
```

#### Cause Analysis

On FusionInsight Manager, the IP address of the Yarn service is in the 192 network segment.

In Yarn logs, the IP address of Spark web UI read by Yarn is http://10.10.169.53:23011, which is in the 10 network segment. The IP addresses in the 192 network segment cannot communicate with those in the 10 network segment. As a result, the Spark web UI fails to be accessed.

#### Procedure

**Step 1** Log in to the client whose IP address is **10.10.169.53**.

**Step 2** Modify the **/etc/hosts** file and change 10.10.169.53 to the IP address of the 192 network segment.

**Step 3** Modify and run the Spark application again. Then, the Spark Web UI is displayed.

----End

### 3.3.14 [Spark WebUI] What Can I Do If an Error Occurs when I Access the Application Page Because the Application Cached by HistoryServer Is Recycled?

#### Question

An error occurs when I access a Spark application page on the HistoryServer page.

Check the HistoryServer logs. The "FileNotFoundException" exception is found. The related logs are as follows:

```
2020-11-22 23:58:03,694 | WARN | [qtp55429210-232] | /history/application_1479662594976_0001/stages/  
stage/ | org.sparkproject.jetty.servlet.ServletHandler.doHandle(ServletHandler.java:628)  
java.io.FileNotFoundException: ${BIGDATA_HOME}/tmp/spark/jobHistoryTemp/  
blockmgr-5f1f6aca-2303-4290-9845-88fa94d78480/09/temp_shuffle_11f82aaf-e226-46dc-  
b1f0-002751557694 (No such file or directory)
```

#### Answer

If a Spark application with a large number of tasks is run on the HistoryServer page, the memory overflows to disk and files with the **temp\_shuffle** prefix are generated.

By default, HistoryServer caches 50 Spark applications (determined by the **spark.history.retainedApplications** configuration item). When the number of Spark applications in the memory exceeds 50, HistoryServer reclaims the first cached Spark application and clears the corresponding **temp\_shuffle** file.

When a user is viewing Spark applications to be recycled, the **temp\_shuffle** file may not be found. As a result, the current page cannot be accessed.

If the preceding problem occurs, use either of the following methods to solve the problem:

- Access the HistoryServer page of the Spark application again. The correct page information is displayed.
- If more than 50 Spark applications need to be accessed at the same time, increase the value of **spark.history.retainedApplications**.

Log in to FusionInsight Manager, choose **Cluster > Services > Spark2x > Configurations**, and click **All Configurations**. In the navigation tree on the left, choose **JobHistory2x > UI**, and set parameters.

**Table 3-1** Parameter description

Parameter	Description	Default Value
spark.history.retainedApplications	Number of Spark applications cached by HistoryServer. When the number of applications to be cached exceeds the value of this parameter, HistoryServer reclaims the first cached Spark application.	50

### 3.3.15 [Spark WebUI] Why Is the Native Page of an Application in Spark2x JobHistory Displayed Incorrectly?

#### Question

Submit a Spark application that contains millions of tasks in a single job. After the application is complete, when I access the native page of the application in JobHistory, the browser waits for a long time before the page is displayed. If the native page cannot be displayed within 10 minutes, "Proxy Error" is displayed.

**Figure 3-3** Example error information

#### Proxy Error

```
The proxy server received an invalid response from an upstream server.  
The proxy server could not handle the request GET /Spark2x/JobHistory2x/77/history/application /1/jobs/  
Reason: Error reading from remote server
```

#### Answer

When you switch to the native page of an application in JobHistory, JobHistory needs to replay the event logs of the application. If the application contains a large number of event logs, the replay takes a long time and the browser waits for a long time.

When the browser accesses the native page in JobHistory, the httpd proxy is required. The timeout interval of the proxy is 10 minutes. Therefore, if JobHistory cannot parse and return the event logs within 10 minutes, httpd returns the "Proxy Error" message to the browser.

#### Solution

The local disk caching function is enabled for JobHistory. When an application is accessed, the parsing result of the event logs of the application is cached to the local disk. When the application is accessed for the second time, the response speed is significantly improved. In this case, you only need to wait for a while and access the original link again.

### 3.3.16 [Spark WebUI] The Spark2x WebUI fails to be accessed using the Internet Explorer

#### Question

When you access the Spark WebUI using Internet Explorer, the access fails occasionally.

#### Symptom

The page fails to be accessed, and the browser cannot display the page, as shown in the following figure.



Turn on TLS 1.0, TLS 1.1, and TLS 1.2 in Advanced settings and try connecting to

#### Cause Analysis

Some earlier versions of Internet Explorer fail to process SSL handshake.

#### Solution

Refresh the page or use Google Chrome to access the page.

### 3.3.17 [Yarn Web UI] Failed to Access the Yarn Web UI

#### Symptom

After the number of nodes in a cluster is expanded to 300, the Yarn web UI cannot be accessed.

#### Possible Causes

When there are a large number of nodes in the cluster, the NodeManager data increases, but the instance memory is not modified. The garbage collection of the ResourceManager process takes a long time, affecting the normal service provision of the ResourceManager process. As a result, an exception occurs when the native Yarn page is accessed.

In this case, it is advised to modify the memory of the instance.

#### Procedure

- Step 1** Log in to FusionInsight Manager and choose **Cluster > Services > Yarn**. On the displayed page, click **Configurations**, click **All Configurations**, and choose **ResourceManager(Role) > System**. Increase the value of the **GC\_OPTS** parameter as required.

 NOTE

The mapping between the number of NodeManager instances in the cluster and the ResourceManager memory size is as follows:

- If the number of NodeManager instances in the cluster reaches 100, it is advised to set the JVM parameter of the ResourceManager instance to **-Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=1G**.
- If the number of NodeManager instances in the cluster reaches 200, it is advised to set the JVM parameter of the ResourceManager instance to **-Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=1G**.
- If the number of NodeManager instances in the cluster reaches 500, it is advised to set the JVM parameter of the ResourceManager instance to **-Xms10G -Xmx10G -XX:NewSize=1G -XX:MaxNewSize=2G**.
- If the number of NodeManager instances in the cluster reaches 1,000, it is advised to set the JVM parameter of the ResourceManager instance to **-Xms20G -Xmx20G -XX:NewSize=1G -XX:MaxNewSize=2G**.

**Step 2** Save the configurations and restart the ResourceManager instance.

**Step 3** Log in to the Yarn web UI again. If the access fails, contact O&M personnel.

----End

# 4 APIs

---

## 4.1 Failed to Call an API to Create a Cluster

### Issue

After the user follows instructions in [Creating a Cluster and Executing a Job](#) to call the cluster creation API `POST /v1.1/{project_id}/run-job-flow`, `500` is returned.

### Symptom

After the cluster creation API `POST /v1.1/{project_id}/run-job-flow` is called, `500` is returned.

### Cause Analysis

The user does not have the permission to create a cluster.

### Procedure

Check whether the user has the permission to create an MRS cluster. For details, see [Permissions Management](#).

# 5 Cluster Management

---

## 5.1 Failed to Reduce Task Nodes

### Symptom

A user fails to reduce the number of task nodes to **0** on details page of the MRS 2.x cluster, and the following information is displayed:

This operation is not allowed because the number of instances of NodeManager will be less than the minimum configuration after scale-in, which may cause data loss.

### Cause Analysis

The NodeManager service of the core node is stopped. If the number of task nodes is changed to **0**, there will be no NodeManager in the cluster and the Yarn service will be unavailable. Therefore, MRS allows the reduction of task nodes only when the number of NodeManagers is greater than or equal to **1**.

### Procedure

**Step 1** Go to the Yarn instance page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console and choose **Components > Yarn > Instances**.

 **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 3.x or later: Log in to FusionInsight Manager, choose **Cluster**, click the name of the target cluster, and choose **Services > Yarn > Instance**.

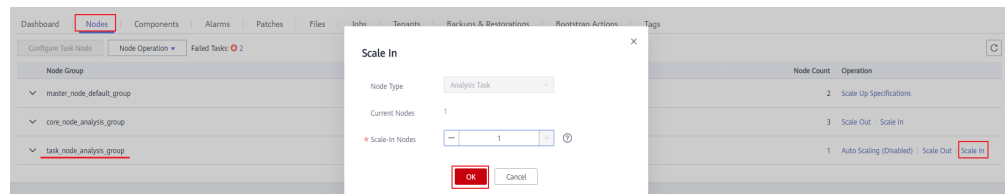
**Step 2** Select the NodeManager instance of the core node, click **More**, and select **Start Instance**.

Role	Host Name	OM IP Address	Business IP Address	Rack	Operating Status
NodeManager	node-core-JWAKw	192.168.0.174	192.168.0.174	/default/rack1b75	Stopped

**Step 3** Reduce the number of task nodes on the cluster details page.

1. Click the cluster name, and select the **Nodes** tab.
2. Locate the row that contains the task node group and click **Scale In** in the **Operation** column.

**Figure 5-1** Removing task nodes



3. Click **OK**. In the displayed dialog box, click **Yes**.

**Step 4** After the scale-in is successful, stop NodeManager of the core node if you do not need it.

----End

## Summary and Suggestions

You are advised not to stop NodeManager of the core node.

## 5.2 OBS Certificate in a Cluster Expired

### Issue

The certificate has expired when a user attempts to access OBS from an MRS cluster.

### Symptom

**ALM-12054 Invalid Certificate File** or **ALM-12055 Certificate File About to Expire** is generated for the MRS cluster, and the certificate that triggers the alarm in the alarm details is the OBS certificate.



**Figure 5-2 Alarm Certificate File About to Expire**

Certificate File About to Expire		12055	Minor	May 9, 2023 00:02:46 G...	OMS	controller	Source=OMS,ServiceN...	Clear	Mask	View Help
Alarm ID:	12055	Alarm Name:	Certificate File About to Expire							
Alarm Severity:	Minor	Source:	OMS							
Generated:	May 9, 2023 00:02:46 GMT+08:00	Cleared:	--							
Object:	controller	Automatic Clearance:	Yes							
Alarm Status:	Uncleared	Alarm Cause:	The certificate file is about to expire.							
Serial Number:	154	Additional Information:	Detail=Certificate availability time < 30 days							
Location:	Source=OMS,ServiceName=controller,RoleName=cacerts/									

**Figure 5-3 Alarm Invalid Certificate File**

Invalid Certificate File		12054	Major	May 8, 2024 18:26:08 G...	OMS	controller	Source=OMS,ServiceN...	Clear	Mask	View Help
Alarm ID:	12054	Alarm Name:	Invalid Certificate File							
Alarm Severity:	Major	Source:	OMS							
Generated:	May 8, 2024 18:26:08 GMT+08:00	Cleared:	May 8, 2023 20:44:37 GMT+08:00							
Object:	controller	Automatic Clearance:	Yes							
Alarm Status:	Automatically Cleared	Alarm Cause:	The certificate file has expired.							
Serial Number:	103	Additional Information:	Detail=Certificate has expired							
Location:	Source=OMS,ServiceName=controller,RoleName=cacerts/ HostName= node-master1@JUL									

## Cause Analysis

The certificate generated by OBS has a validity period. After the validity period expires, the certificate file becomes invalid, and an alarm is generated.

## Procedure

**Step 1** Query the OBS certificate information of the MRS cluster.

Log in to the active OMS node of the MRS cluster as user **root** and run the following command to check whether an OBS certificate exists:

```
keytool -v -list -keystore ${JAVA_HOME}/jre/lib/security/cacerts -protected 2> /dev/null|grep -E "Alias name*|Valid from*" | grep obs
```

An OBS certificate exists if information similar to the following is returned:

```
Alias name: obs.example.com
```

- If no certificate exists, no further action is required. Wait until the alarm is cleared.
- If a certificate exists, go to **Step 2**.

### NOTE

- `${java_home}` indicates the JDK directory of the cluster. In MRS 3.x, replace it with `/opt/Bigdata/common/runtime0/jdk1.8*`. In versions earlier than MRS 3.x, replace it with `/opt/Bigdata/jdk`.
- In MRS 3.x, if the certificate expiration alarm persists even after you perform the operations provided in this section, replace `${JAVA_HOME}` with `Client installation directory/JDK/jdk` and perform the operations again.

**Step 2** Delete the OBS certificate.

On the active OMS node, run the following commands to delete the OBS certificate queried in **Step 1**:

```
obs_url=$(keytool -v -list -keystore ${JAVA_HOME}/jre/lib/security/cacerts -  
protected 2> /dev/null|grep -E "Alias name*|Valid from*" | grep obs | cut -d ':'  
-f 2 | awk '$1=$1')
```

```
jdk_cacert="${JAVA_HOME}/jre/lib/security/cacerts"
```

```
keytool -delete -alias ${obs_url} -keystore ${jdk_cacert} -storepass changeit
```

- Step 3** Run the following command to check that the OBS certificate does not exist. If the certificate still exists, go to [Step 2](#).

```
keytool -v -list -keystore ${JAVA_HOME}/jre/lib/security/cacerts -protected  
2> /dev/null|grep -E "Alias name*|Valid from*" | grep obs
```

```
----End
```

## 5.3 Replacing a Disk in an MRS Cluster (Applicable to 2.x and Earlier)

### Issue

A disk is not accessible.

### Symptom

A user created an MRS cluster with local disks. A disk of a core node in this cluster is damaged, resulting in file read failures.

### Cause Analysis

The disk hardware is faulty.

### Procedure

#### NOTE

This procedure is applicable to analysis clusters earlier than MRS 3.x. If you need to replace disks for a streaming cluster or hybrid cluster, contact Huawei Cloud technical support.

- Step 1** Log in to [MRS Manager](#).

- Step 2** Choose **Hosts**, click the name of the target host, click **RegionServer** in the **Roles** list, click **More**, and select **Decommission**.

- Step 3** Choose **Hosts**, click the name of the target host, click **DataNode** in the **Roles** list, click **More**, and select **Decommission**.

- Step 4** Choose **Hosts**, click the name of the target host, click **NodeManager** in the **Roles** list, click **More**, and select **Decommission**.

#### NOTE

If this host still runs other instances, perform the similar operation to decommission the instances.

- Step 5** Run the `vim /etc/fstab` command to comment out the mount point of the faulty disk.

**Figure 5-4** Commenting out the mount point of the faulty disk

```
[root@node-ana-coreqexX0001 ~]# vim /etc/fstab
devpts /dev/pts          devpts mode=0620,gid=5 0 0
proc   /proc                 proc   defaults                0 0
sysfs  /sys                  sysfs  noauto                  0 0
debugfs /sys/kernel/debug    debugfs noauto                  0 0
tmpfs  /run                  tmpfs  noauto                  0 0
/dev/disk/by-label/ROOT / ext4 defaults,noatime 1 1
UUID=0f871b41-61e0-4f7f-af54-a03a1bfb3753 /srv/BigData/hadoop/data1 ext4 defaults,noatime,nodiratime 1 0
```

- Step 6** If the old disk is still accessible, migrate user data on the old disk (for example, `/srv/BigData/hadoop/data1/`).
- Step 7** Log in to the MRS console.
- Step 8** On the cluster details page, click the **Nodes** tab.
- Step 9** Click the node whose disk is to be replaced to go to the ECS console. Click **Stop** to stop the node.
- Step 10** Contact Huawei Cloud technical support to replace the disk in the background.
- Step 11** On the ECS console, click **Start** to start the node where the disk has been replaced.
- Step 12** Run the `fdisk -l` command to view the new disk.
- Step 13** Run the `cat /etc/fstab` command to obtain the drive letter.

**Figure 5-5** Obtaining the drive letter

```
[omm@node-master1dGom ~]$ cat /etc/fstab
#
# /etc/fstab
# Created by anaconda on Wed Feb 27 06:58:49 2019
#
# Accessible filesystems, by reference, are maintained under '/dev/disk'
# See man pages fstab(5), findfs(8), mount(8) and/or blkid(8) for more info
#
UUID=b13ee9c8-0ef0-4159-9b90-fc47bde0d464 / ext4 defaults,noatime 1 1
UUID=029408e0-71a6-4f73-b817-42d7049b7595 /srv/BigData ext4 defaults,noatime,nodiratime 1 0
UUID=f9cb8844-dabf-4a69-aff4-587de2fc4d7c /srv/BigData1 ext4 defaults,noatime,nodiratime 1 0
UUID=876e73be-1f80-4466-92b7-01d7c68bbb1b /srv/BigData2 ext4 defaults,noatime,nodiratime 1 0
UUID=0d5fce7f-afd0-420a-b1bb-e5500a1851cd /srv/BigData3 ext4 defaults,noatime,nodiratime 1 0
```

- Step 14** Use the corresponding drive letter to format the new disk.  
Example: `mkfs.ext4 /dev/sdh`
- Step 15** Run the following command to attach the new disk.  
`mount New disk Mount point`  
Example: `mount /dev/sdh /srv/BigData/hadoop/data1`
- Step 16** Run the following command to grant the `omm` user permission to the new disk:  
`chown omm:wheel Mount point`  
Example: `chown -R omm:wheel /srv/BigData/hadoop/data1`
- Step 17** Add the UUID of the new disk to the `fstab` file.
1. Run the `blkid` command to check the UUID of the new disk.

```
[root@node-ana-corekpoT0003 ~]# blkid
/dev/uda1: LABEL="ROOT" UUID="2aa97872-11ec-422e-9513-0f28b925ad5e" TYPE="ext4"
/dev/udb: UUID="e5f652c3-f9af-427f-89da-f2545618688d" TYPE="ext4"
[root@node-ana-corekpoT0003 ~]#
```

2. Open the `/etc/fstab` file and add the following information:  
`UUID=New disk UUID /srv/BigData/hadoop/data1 ext4 defaults,noatime,nodiratime 1 0`

**Step 18** (Optional) Create a log directory.

```
mkdir -p /srv/BigData/Bigdata
chown omm:ficommon /srv/BigData/Bigdata
chmod 770 /srv/BigData/Bigdata
```

 **NOTE**

Run the following command to check whether symbolic links to **Bigdata** logs exist. If yes, skip this step.

```
ll /var/log
```

**Step 19** Log in to **MRS Manager**.

**Step 20** Choose **Hosts**, click the name of the target host, click **RegionServer** in the **Roles** list, click **More**, and select **Recommission**.

**Step 21** Choose **Hosts**, click the name of the target host, click **DataNode** in the **Roles** list, click **More**, and select **Recommission**.

**Step 22** Choose **Hosts**, click the name of the target host, click **NodeManager** in the **Roles** list, click **More**, and select **Recommission**.

 **NOTE**

If this host still runs other instances, perform the similar operation to recommission the instances.

**Step 23** Choose **Services > HDFS**. In the **HDFS Summary** area on the **Service Status** page, check whether **Missing Blocks** is **0**.

- If **Missing Blocks** is **0**, no further action is required.
- If **Missing Blocks** is not **0**, contact Huawei Cloud technical support.

----End

## 5.4 Replacing a Disk in an MRS Cluster (Applicable to 3.x)

### Issue

A disk is not accessible.

### Symptom

A user created an MRS cluster with local disks. A disk of a core node in this cluster is damaged, resulting in file read failures.

### Cause Analysis

The disk hardware is faulty.

## Procedure

### NOTE

This procedure is applicable to troubleshooting disk hardware faults of core and task nodes in MRS clusters using local disks (ECSs of D, I, IR, and KI series).

Kafka does not support disk replacement. If the node that stores Kafka data is faulty, contact Huawei Cloud technical support.

**Step 1** Log in to [FusionInsight Manager](#).

**Step 2** Choose **Hosts** and click the name of the faulty host. In the **Instance** area, click **DataNode**. Then on the page that is displayed, click **More** and select **Decommission**.

### NOTE

- If this host accommodates DataNode, NodeManager, RegionServer, and ClickHouseServer instances, decommission these instances by referring to this step.
- In versions later than MRS 3.1.2, the ClickHouseServer role instance can be decommissioned.

**Step 3** Choose **Hosts**, select the faulty host, click **More**, and select **Stop All Instances**.

**Step 4** Run the `vim /etc/fstab` command to comment out the mount point of the faulty disk.

**Figure 5-6** Commenting out the mount point of the faulty disk

```
[root@node-ana-coreXZYb0001 ~]# vim /etc/fstab
#
# /etc/fstab
# Created by anaconda on Sat Feb 27 07:10:42 2021
#
# Accessible filesystems, by reference, are maintained under '/dev/disk'
# See man pages fstab(5), findfs(8), mount(8) and/or blkid(8) for more info
#
UUID=c89eca08-5da4-43de-add0-4bb58e820d78 / ext4 defaults,errors=panic,noatime 1 1
UUID=4b16f96b-6d16-4d8e-9517-9f63423f9f6e /tmp ext4 defaults,noatime,nodiratime,errors=panic 1 0
UUID=e539a0fd-a639-41dc-aa88-5f6c0e4bb7b3 /var ext4 defaults,noatime,nodiratime,errors=panic 1 0
UUID=51ba7a26-67de-4762-8bea-85fc004065c2 /srv/BigData ext4 defaults,noatime,nodiratime 1 0
UUID=03ba5f78-d188-4e6b-b640-1915b858183a /var/log ext4 defaults,noatime,nodiratime,errors=panic 1 0
# UUID=91c84554-22eb-4130-a7a1-5ceaf03c8c06 /srv/BigData/data1 ext4 defaults,noatime,nodiratime,nodev 1 0
```

**Step 5** If the old disk is still accessible, migrate user data on the old disk (for example, `/srv/BigData/data1/`).

**cp -r Mount point of the old disk Temporary data storage directory**

Example: `cp -r /srv/BigData/data1 /tmp/`

**Step 6** Log in to the MRS console.

**Step 7** On the cluster details page, click the **Nodes** tab.

**Step 8** Click the node whose disk is to be replaced to go to the ECS console. Click **Stop** to stop the node.

**Step 9** Contact Huawei Cloud technical support to replace the disk in the background.

**Step 10** On the ECS console, click **Start** to start the node where the disk has been replaced.

**Step 11** Initialize the Linux data disk.

For details, see [Initializing EVS Data Disks](#).

**Step 12** Run the **lsblk** command to view information about the new disk partition.

**Figure 5-7** Viewing the new disk partition

```
[root@ecs-fcq ~]# lsblk
NAME        MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda          8:0    0  1.7T  0 disk
sdb          8:16   0  1.7T  0 disk
sdc          8:32   0  1.7T  0 disk
└─sdc1       8:33   0  1.7T  0 part
sdd          8:48   0  1.7T  0 disk
└─sdd1       8:49   0  1.7T  0 part
```

**Step 13** Run the **df -TH** command to obtain the file system type.

**Figure 5-8** Obtaining the file system type

```
[root@node-ana-corewQaI0001 ~]# df -TH
Filesystem      Type      Size  Used Avail Use% Mounted on
/dev/vda1       ext4      233G  44G  179G  20% /
devtmpfs        devtmpfs  34G   0    34G   0% /dev
tmpfs           tmpfs     34G   0    34G   0% /dev/shm
tmpfs           tmpfs     34G   9.3M 34G   1% /run
tmpfs           tmpfs     34G   0    34G   0% /sys/fs/cgroup
/dev/vda5       ext4      11G   40M  10G   1% /tmp
/dev/vda7       ext4      64G   152M 60G   1% /srv/BigData
/dev/vda6       ext4      11G   1.2G 8.9G  12% /var
/dev/vda8       ext4      190G  211M 180G   1% /var/log
/dev/sdc1       ext4      1.8T  1.4G 1.8T   1% /srv/BigData/data2
tmpfs           tmpfs     6.8G   0    6.8G   0% /run/user/2000
tmpfs           tmpfs     6.8G   0    6.8G   0% /run/user/0
[root@node-ana-corewQaI0001 ~]#
```

**Step 14** Format the new disk partition based on the obtained file system type.

Example: **mkfs.ext4 /dev/sdd1**

**Step 15** Run the following command to mount the new disk:

**mount** *New disk Mount point*

Example: **mount /dev/sdd1 /srv/BigData/data1**

**NOTE**

If the disk cannot be mounted, run the following command to reload the configuration and mount it again:

**systemctl daemon-reload**

**Step 16** Run the following command to grant the **omm** user permission to the new disk:

**chown omm:wheel** *Mount point*

Example: **chown -R omm:wheel /srv/BigData/data1**

**Step 17** Migrate user data from the old disk (for example, **/srv/BigData/data1/**) to the new disk.

**cp -r** *Temporary data storage directory Mount point of the new disk*

Example: `cp -r /tmp/data1/* /srv/BigData/data1/`

**Step 18** Add the UUID of the new disk to the `fstab` file.

1. Run the `blkid` command to check the UUID of the new disk.

```
root@node-ana-core10001 ~# blkid
/dev/sda6: UUID="e539a0fd-a639-41dc-aa88-5f4c8e4bb7b3" TYPE="ext4"
/dev/sda1: UUID="c89eca08-5da4-43de-ada0-4bb58e820d78" TYPE="ext4"
/dev/sda5: UUID="4b16f96b-6d16-4d8e-9517-9f63423f9f6e" TYPE="ext4"
/dev/sda7: UUID="51ba7a26-67de-4762-bbca-85fc004065c2" TYPE="ext4"
/dev/sda8: UUID="03ba5f78-d188-4e6b-b648-19153858183a" TYPE="ext4"
/dev/sda1: UUID="02a89811-ac36-4148-abad-e5ef935e54e8" TYPE="ext4" PARTLABEL="logical1" PARTUUID="1bd64663-42e1-4bdf-9ece-4b5b793cf799"
/dev/sdc1: UUID="578ccafe-4585-462a-a358-e12488969d7f" TYPE="ext4" PARTLABEL="logical1" PARTUUID="ac389415-3294-47c4-b089-ae39fc72f62e"
/dev/sdd1: UUID="7f377c8b-e1b9-423e-b7d2-a60e1d58c3eb" TYPE="ext4" PARTLABEL="logical1" PARTUUID="7f8254ea-386c-46ae-b358-8e3845de5128"
/dev/sdb1: UUID="67133dc9-da39-4561-9353-602257347cc1" TYPE="ext4" PARTLABEL="logical1" PARTUUID="2004ff81-e343-4f41-bfe8-889b4bd38968"
root@node-ana-core10001 ~#
```

2. Open the `/etc/fstab` file and add the following information:

```
UUID=UUID of the new disk /srv/BigData/data1 ext4 defaults,noatime,nodiratime,nodew 1 0
```

**Step 19** Log in to [FusionInsight Manager](#).

**Step 20** Choose **Hosts** and click the name of the host to be recommissioned. In the **Instance** area, click **DataNode**. Then on the page that is displayed, click **More** and select **Recommission**.

#### NOTE

- If this host accommodates `DataNode`, `NodeManager`, `RegionServer`, and `ClickHouseServer` instances, recommission these instances by referring to this step.
- In versions later than MRS 3.1.2, the `ClickHouseServer` role instance can be recommissioned.

**Step 21** Choose **Hosts**, select the faulty host, click **More**, and select **Start All Instances**.

**Step 22** Choose **Cluster > HDFS**. In the **Basic Information** area on the **Dashboard** page, check whether **Missing Blocks** is **0**.

- If **Missing Blocks** is **0**, no further action is required.
- If **Missing Blocks** is not **0**, contact Huawei Cloud technical support.

----End

## 5.5 Failed to Execute an MRS Backup Task

### Symptom

MRS backup tasks always fail.

### Cause Analysis

The backup directory points to the system disk using a symbolic link. If the system disk is full, the backup task fails.

### Procedure

**Step 1** Check whether the backup directory points to the system disk using a symbolic link.

1. Log in to the active and standby master nodes in the cluster as user `root`.

2. Run the **df -h** command to check the storage usage of the system disk.
3. Run the **ll /srv/BigData/LocalBackup** command to check whether the backup directory points to **/opt/Bigdata/LocalBackup** using a symbolic link.

Check whether the backup file points to the system disk using a symbolic link and whether the system disk has sufficient space. If the symbolic link points to the system disk and the system disk space is insufficient, go to [Step 2](#). If the symbolic link is not used, the failure is not caused by insufficient system disk space. Contact technical support for troubleshooting.

**Step 2** Move historical backup data to a new directory on the data disk.

1. Log in to the master node as user **root**.
2. Run the **su - omm** command to switch to user **omm**.
3. Run the **rm -rf /srv/BigData/LocalBackup** command to delete the symbolic link of the backup directory.
4. Run the **mkdir -p /srv/BigData/LocalBackup** command to create a backup directory.
5. Run the **mv /opt/Bigdata/LocalBackup/\* /srv/BigData/LocalBackup/** command to move the historical backup data to the new directory.

**Step 3** Execute the backup task again.

----End

## 5.6 Inconsistency Between df and du Command Output on the Core Node

### Symptom

After the **df** and **du** commands are executed, the values of the core node capacity displayed are different.

The disk usage of the **/srv/BigData/hadoop/data1/** directory queried by running the **df -h** command differs greatly from that queried by running the **du -sh /srv/BigData/hadoop/data1/** command. The difference is greater than 10 GB.

### Cause Analysis

The **lsdf |grep deleted** command output indicates that a large number of log files in the directory are in the deleted state.

When some Spark tasks are running for a long time, some containers in the tasks keep running and logs are continuously generated. When printing logs, the executor of Spark uses the log4j log scrolling function to output logs to the **stdout** file. The container also monitors this file. As a result, the file is monitored by two processes at the same time. When one process scrolls according to the configuration, the earliest log file is deleted, but the other process still occupies the file handle. As a result, a file in the deleted state is generated.

### Procedure

Change the output directory name for executor logs of Spark.



**Step 1** Open the log configuration file. By default, the configuration file is located in *<Client installation directory>/Spark/spark/conf/log4j-executor.properties*.

**Step 2** Change the name of the log output file.

For example:

```
log4j.appender.sparklog.File = ${spark.yarn.app.container.log.dir}/stdout
```

is changed to

```
log4j.appender.sparklog.File = ${spark.yarn.app.container.log.dir}/stdout.log
```

**Step 3** Save the configuration and exit.

**Step 4** Submit the task again.

----End

## 5.7 Disassociating a Subnet from a Network ACL

### Scenarios

You can disassociate a subnet from a network ACL when necessary.

### Procedure

**Step 1** Log in to the management console.

**Step 2** In the **Service List**, choose **Networking > Virtual Private Cloud**.

**Step 3** In the navigation pane on the left, choose **Access Control > Network ACLs**.

**Step 4** On the **Network ACLs** page in the right pane, click the target network ACL name to switch to its details page.

**Step 5** On the displayed page, click the **Associated Subnets** tab.

**Step 6** On the **Associated Subnets** page, locate the target network ACL and click **Disassociate** in the **Operation** column.

**Step 7** Click **OK**.

----End

## 5.8 MRS Cluster Becomes Abnormal After the Hostname of a Node Is Changed

### Symptom

After the hostname of a node is changed, the MRS cluster becomes abnormal.

### Cause Analysis

Changing the hostname causes compatibility problems and faults.

## Procedure

- Step 1** Log in to any node in the cluster as user **root**.
- Step 2** Run the **cat /etc/hosts** command on the node to check the value of **hostname** of each node and set the **newhostname** variable based on the value.
- Step 3** Run the **sudo hostnamectl set-hostname \${newhostname}** command on the node where **hostname** is changed to restore the correct hostname.

 **NOTE**

**\${newhostname}**: new value of **hostname**

- Step 4** Log in to the node where **hostname** is changed, and check whether the new hostname takes effect.

----End

## 5.9 Processes Are Terminated Unexpectedly

### Symptom

A DataNode is restarted unexpectedly, but no manual restart operation is performed for the DataNode.

### Cause Analysis

Possible causes:

- **OOM of the Java process is terminated.**

In general, the OMM Killer is configured for Java processes to detect and kill OOM. The OOM log is printed in the out log. In this case, you can view the run log (for example, the DataNode's log path is **/var/log/Bigdata/hdfs/dn/hadoop-omm-datanode-*hostname*.log**) to check whether OutOfMemory is printed.

- **The process is terminated by another process or manually terminated.**

Check the DataNode run log file **/var/log/Bigdata/hdfs/dn/hadoop-omm-datanode-*hostname*.log**. It is found that the health check fails after "RECEIVED SIGNAL 15" is received.

In the following example, the DataNode is terminated at 11:04:48 and then started at 11:06:52 two minutes later.

```
2018-12-06 11:04:48,433 | ERROR | SIGTERM handler | RECEIVED SIGNAL 15: SIGTERM |
LogAdapter.java:69
2018-12-06 11:04:48,436 | INFO | Thread-1 | SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down DataNode at 192-168-235-85/192.168.235.85
*****/ LogAdapter.java:45
2018-12-06 11:06:52,744 | INFO | main | STARTUP_MSG:
```

According to the logs, DataNode was closed and then the health check reported the exception. After 2 minutes, NodeAgent started the DataNode process.

## Procedure

Add the rule for recording the **kill** command in the audit log of the OS. The process that delivers the command will be recorded in the audit log.

### Operation impact

- Printing audit logs affects operating system performance. However, analysis result shows that the impact is less than 1%.
- Printing audit log occupies some disk space. The logs to be printed are within megabytes. By default, the aging mechanism and the mechanism for checking the remaining disk space are configured. Therefore, the disk space will not be used up.

### Locating Method

Perform the following operations on nodes that may restart the DataNode process:

- Step 1** Log in to the node as the **root** user and run the **service auditd status** command to check the service status.

```
Checking for service auditd running
```

If the service is not started, run the **service auditd restart** command to restart the service. The command execution takes less than 1 second and has no impact on the system.

```
Shutting down auditd done  
Starting auditd done
```

- Step 2** The audit rule of the **kill** command is temporarily added to audit logs.

Add an audit rule:

```
auditctl -a exit,always -F arch=b64 -S kill -S tkill -S tkill -F a1!=0 -k process_killed
```

View the rule:

```
auditctl -l
```

- Step 3** If a process is terminated due to an exception, run the **ausearch -k process\_killed** command to query the termination history.

```
[root@aaaa ~]# ausearch -k process_killed  
----  
time->Fri Jul 8 15:43:44 2016  
type=CONFIG_CHANGE msg=audit(1467963824.969:48328): auid=0 ses=3514 subj=unconfined_u:system_r:auditctl_t:s0 op="add rule" key="process_killed" list=4 res=1  
----  
time->Fri Jul 8 15:43:50 2016  
type=OBJ_PID msg=audit(1467963830.034:48329): opid=21601 cauid=0 coid=0 cses=3965 obj=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 ocomm="diskmgt0"  
type=SYSCALL msg=audit(1467963830.034:48329): arch=0000003e syscall=62 success=yes exit=0 a0=5461 a1=0 a2=0 a3=5461 items=0 ppid=6919 pid=14173 auid=0 uid=0 gid=0 euid=0 egid=0  
=0 euid=0 fsuid=0 tty=pts1 ses=3514 comm="bash" exe="/bin/bash" subj=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 key="process_killed"
```

### NOTE

**a0** is the PID (hexadecimal) of the process that is terminated, and **a1** is the semaphore of the **kill** command.

----End

### Verification

**Step 1** Restart an instance of the node on MRS Manager, for example, DataNode.

**Step 2** Run the **ausearch -k process\_killed** command to check whether logs are printed.

The following is an example of the **ausearch -k process\_killed |grep ".sh"** command. The command output indicates that the **hdfs-daemon-ada\*** script closed the DataNode process.

```

[root@5-148-0 Bigdata]# ausearch -k process_killed | grep ".sh"
type=SYSCALL msg=audit(1481027379:225:2269996): arch=0000003e syscall=62 success=yes exit=0 a0=78dc a1=f a2=0 a3=78dc item=0 ppid=2873 pid=2888 uid=2000 gid=10 euid=2000 suid=2000 fsuid=2000 egid=10 sgid=10 fsgid=10 tty=ln
one) ses=10 comm="hdfs-daemon-ada" exe="/bin/bash" subj=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 key="process_killed"
type=SYSCALL msg=audit(1481027379:225:2269996): arch=0000003e syscall=60 success=yes exit=0 a0=78dc a1=0 a2=0 a3=ffffa2a000 item=0 ppid=2873 pid=2888 uid=2000 gid=10 euid=2000 suid=2000 fsuid=2000 egid=10 sgid=10 fsgid=1
0 tty=lnone) ses=10 comm="hdfs-daemon-ada" exe="/bin/bash" subj=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 key="process_killed"
type=SYSCALL msg=audit(1481027379:225:2269996): arch=0000003e syscall=62 success=no exit=-9 a0=78dc a1=0 a2=0 a3=78dc item=0 ppid=2873 pid=2888 uid=2000 gid=10 euid=2000 suid=2000 fsuid=2000 egid=10 sgid=10 fsgid=10 tty=ln
one) ses=10 comm="hdfs-daemon-ada" exe="/bin/bash" subj=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 key="process_killed"
[root@5-148-0 Bigdata]#

```

----End

Stop auditing the **kill** command.

**Step 1** Run the **service auditd restart** command. The temporarily added **kill** command audit logs are cleared automatically.

**Step 2** Run the **auditctl -l** command. If no related information is displayed, the rule is cleared successfully.

----End

## 5.10 Failed to Configure Cross-Cluster Mutual Trust for MRS

### Symptom

The cross-cluster mutual trust relationship cannot be established between a cluster earlier than MRS 1.8.2 and a cluster later than MRS 1.8.2.

### Cause Analysis

After cross-cluster mutual trust is configured, users **krbtgt/Local cluster domain name@External cluster domain name** and **krbtgt/External cluster domain name@Local cluster domain name** are added to both clusters. The default passwords for the users of the two clusters are different. As a result, cross-cluster mutual trust fails to be configured.

### Procedure

- **Scenario without mutual trust being configured:**
  - a. Before configuring the mutual trust, log in to the Master node in the cluster of MRS 1.8.2 or later.
  - b. Change the value of **local cross\_realm\_account\_pwd="{DEFAULT\_CROSS\_REALM\_PWD}"** in the **add\_cross\_realm\_princ** method of the **/opt/Bigdata/om-0.0.1/sbin/addRealm.sh** script on all master nodes to **local cross\_realm\_account\_pwd="{DEFAULT\_PWD}"** (in line 1001 of the script).

#### NOTE

Perform steps from **a** to **b** on all master nodes in the cluster of MRS 1.8.2 or later.

- c. Then, configure cross-cluster mutual trust by referring to [Configuring Cross-Cluster Mutual Trust Relationships](#).
- d. Check whether the mutual trust relationship is established.
  - If yes, the configuration is complete.
  - If the relationship fails to be established, refresh the client configuration and check whether the trust relationship is established. If the problem persists, contact O&M personnel.
- **Scenario with mutual trust being configured**
  - a. Log in to the master node in the cluster of MRS 1.8.2 or later.
  - b. Run the `/home/omm/kerberos/bin/kadmin -p kadmin/admin` command and enter the password of the Kerberos client.
  - c. Run the `listprincs` command and press **Enter** to query user information.

```

Last login: Thu Nov 14 12:53:41 2019 from 10.190.111.196

[root@node-master1EnmC ~]# /home/omm/kerberos/bin/kadmin -p kadmin/admin
Authenticating as principal kadmin/admin with password.
Password for kadmin/admin@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM:
kadmin: listprincs
K/M@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
krbtgt/8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
kadmin/admin@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
kadmin/changepw@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
kadmin/history@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
backup/manager@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
admin@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
oms/manager@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
executor@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
knox@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
hdfs@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
hdfs/hadoop.8af4b8e5_a6fe_4b60_9bdd_efd84c8e024c.com@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
zookeeper/hadoop.8af4b8e5_a6fe_4b60_9bdd_efd84c8e024c.com@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
zkcli/hadoop.8af4b8e5_a6fe_4b60_9bdd_efd84c8e024c.com@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
mapred@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
mapred/hadoop.8af4b8e5_a6fe_4b60_9bdd_efd84c8e024c.com@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
mr_zk/hadoop.8af4b8e5_a6fe_4b60_9bdd_efd84c8e024c.com@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
check_ker_M@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
HTTP/node-ana-coreoeyi@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
HTTP/node-master1enmc@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
spark_zk/hadoop.8af4b8e5_a6fe_4b60_9bdd_efd84c8e024c.com@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
krbtgt/8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM@2A30110C_BBFA_468C_A802_4808C6B10B2A.COM
krbtgt/2A30110C_BBFA_468C_A802_4808C6B10B2A.COM@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
zlu_jj@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
test_jj@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
krbtgt/8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM@1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM
krbtgt/1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
kadmin:
    
```

- d. Run the `delprinc` command to delete users `krbtgt/Local cluster domain name@External cluster domain name` and `krbtgt/External cluster domain name@Local cluster domain name`.

```

kadmin: delprinc krbtgt/8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM@1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM
Are you sure you want to delete the principal "krbtgt/8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM@1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM"? (yes/no): yes
Principal "krbtgt/8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM@1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM" deleted.
Make sure that you have removed this principal from all ACLs before reusing.
kadmin: delprinc krbtgt/1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM
Are you sure you want to delete the principal "krbtgt/1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM"? (yes/no): yes
Principal "krbtgt/1353EEF8_DBFB_4ECA_8669_BC0B770A8AB.COM@8AF4B8E5_A6FE_4B60_9BDD_EFD84C8E024C.COM" deleted.
Make sure that you have removed this principal from all ACLs before reusing.
kadmin:
    
```

- e. Run the `quit` command to exit the Kerberos client.
- f. Change the value of `local cross_realm_account_pwd="$ {DEFAULT_CROSS_REALM_PWD}"` in the `add_cross_realm_princ` method of the `/opt/Bigdata/om-0.0.1/sbin/addRealm.sh` script on the master nodes to `local cross_realm_account_pwd="$ {DEFAULT_PWD}"` (in line 1001 of the script).
- g. Log in to MRS Manager, and choose **Services**.
- h. Click **More** and select **Synchronize Configuration**.
- i. In the dialog box displayed, select **Restart the service or instance whose configuration has expired** and click **OK**.

During configuration synchronization, the `addRealm.sh` script is invoked to add the `krbtgt` user.

**NOTE**

Perform steps from **a** to **i** on all master nodes in the cluster of MRS 1.8.2 or later.

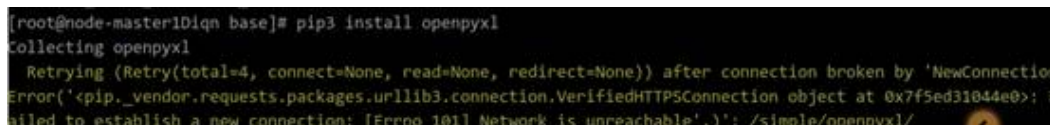
- j. Check whether the mutual trust is established. If it still fails, contact O&M personnel.

## 5.11 Network Is Unreachable When Python Is Installed on an MRS Cluster Node Using Pip3

### Symptom

When a user runs the **pip3 install** command to install Python, an error message is displayed, indicating that the network is unreachable.

For details, see the following figure.



```
[root@node-master1D1qn base]# pip3 install openpyxl
Collecting openpyxl
  Retrying (Retry(total=4, connect=None, read=None, redirect=None)) after connection broken by 'NewConnectionError(<pip._vendor.requests.packages.urllib3.connection.VerifiedHTTPSConnection object at 0x7f5ed31844e0>; Failed to establish a new connection: [Errno 101] Network is unreachable',)': /simple/openpyxl/
```

### Cause Analysis

The user does not bind an EIP to the master node.

### Procedure

- Step 1** Log in to the MRS management console.
  - Step 2** Choose **Active Clusters**, select the faulty cluster, and click its name to check the **Basic Information** on the **Dashboard** tab page.
  - Step 3** On the **Nodes** tab, click the name of a master node in the master node group to log in to the ECS console.
  - Step 4** Click the **EIPs** tab and click **Bind EIP** to bind an EIP to the ECS.
  - Step 5** Log in to the master node and run the **pip3 install** command to install Python.
- End

## 5.12 Connecting the Open-Source confluent-kafka-go to an MRS Security Cluster

### Issue

The user does not know how to connect the Open-Source confluent-kafka-go to an MRS security cluster.

## Symptom

The open-source confluent-kafka-go fails to connect to the security cluster of MRS.

## Cause Analysis

By default, the librdkafka library on which confluent-kafka-go depends uses the hostname of the broker as a part of the server principle. As a result, the authentication fails.

## Procedure

The procedure for modifying the librdkafka is as follows:

1. The librdkafka source code address: <https://github.com/edenhill/librdkafka>
2. Add the **sasl.kerberos.service.name** configuration item to the **src/rdkafka\_conf.c** file.

```
"Kerberos principal name that Kafka runs as.", .sdef = "kafka" }, { _RK_GLOBAL,
"sasl.kerberos.principal", _RK_C_STR, _RK(sasl.principal), "This client's Kerberos
principal name.", .sdef = "kafkaclient" }, + { _RK_GLOBAL, "sasl.kerberos.domain.name",
_RK_C_STR, + _RK(sasl.domain_name), + "This cluster's Kerberos domain name.",
+ .sdef = "hadoop.hadoop.com" }, #ifndef _MSC_VER { _RK_GLOBAL,
"sasl.kerberos.kinit.cmd", _RK_C_STR, _RK(sasl.kinit_cmd), "Full kerberos kinit
command string, %{config.prop.name} is replaced " "by corresponding config object value, %
{broker.name} returns the " "broker's hostname.", - .sdef = "kinit -S \"%
{sasl.kerberos.service.name}/%{broker.name}\" " + .sdef = "kinit -S \"%
{sasl.kerberos.service.name}/%{sasl.kerberos.domain.name}\" " "-k -t \"%{sasl.kerberos.keytab}
\" \"%{sasl.kerberos.principal}\" " }, { _RK_GLOBAL, "sasl.kerberos.keytab", _RK_C_STR,
_RK(sasl.keytab), "Path to Kerberos keytab file. Uses system default if not set."
***NOTE**: This is not automatically used but must be added to the " "template in
sasl.kerberos.kinit.cmd as "
```

3. Add the **domain\_name** field to the **src/rdkafka\_conf.h** file.

```
--- src\rdkafka_conf.h 2017-10-17 11:20:56.000000000 +0800 +++ src\rdkafka_conf.h
2017-10-25 16:26:34.000000000 +0800 @@ -118,12 +118,13 @@ struct { const
struct rd_kafka_sasl_provider *provider; char *principal; char
*mechanisms; char *service_name; + char *domain_name;
char *kinit_cmd; char *keytab; int relogin_min_time; char
*username; char *password; #if WITH_SASL_SCRAM
```

4. Replace **hostname** with **domainName** in the **src/rdkafka\_sasl\_cyrus.c** file.

```
--- src\rdkafka_sasl.c 2017-10-17 11:20:56.000000000 +0800 +++ src\rdkafka_sasl.c
2017-10-25 16:09:38.000000000 +0800 @@ -192,13 +192,14 @@ rk-
>rk_conf.sasl.mechanisms, rk->rk_conf.api_version_request ?
"" : " : try api.version.request=true"); return -1; } -
rd_strdupa(&hostname, rktrans->rktrans_rkb->rkb_nodename); + //rd_strdupa(&hostname,
rktrans->rktrans_rkb->rkb_nodename); + rd_strdupa(&hostname, rk-
>rk_conf.sasl.domain_name); if ((t = strchr(hostname, ':')) *t = '\0'; /* remove
":port" */
```

5. Recompile librdkafka (ensure that libsasl2-dev has been installed). For details, see <https://github.com/edenhill/librdkafka/tree/v0.11.1>.

**./configure make make install**

6. Add the following configuration items when using the client:

```
"security.protocol": "SASL_PLAINTEXT",
"sasl.kerberos.service.name": "kafka",
"sasl.kerberos.keytab": "/opt/nemon/user.keytab",
"sasl.kerberos.principal": "nemon@HADOOP.COM",
"sasl.kerberos.domain.name": "hadoop.hadoop.com",
```

 NOTE

MRS 2.1.x and earlier versions:

- **sas.l.kerberos.keytab**: On MRS Manager, choose **System > Manage User**. Locate the row that contains the target user, choose **More > Download authentication credential**. Save the file and decompress it to extract the **user.keytab** file.
- **sas.l.kerberos.principal**: Enter the actual user name.
- **sas.l.kerberos.domain.name**: The domain naming rule is **hadoop.toLowerCase(realm)**: If the cluster domain name (**default\_realm**) is **example.com**, the value of **domain** is **hadoop.example.com**. On MRS Manager, choose **Services > KrbServer > Service Configuration > All**, and search for and view the value of **default\_realm**.

For MRS 3.x or later:

- **sas.l.kerberos.keytab**: On FusionInsight Manager, choose **System > Permission > User**. Locate the row that contains the target user, click **More**, and select **Download authentication credential** in the **Operation** column. Save the file and decompress it to extract the **user.keytab** file.
- **sas.l.kerberos.principal**: Enter the actual user name.
- **sas.l.kerberos.domain.name**: The domain naming rule is **hadoop.toLowerCase(realm)**: If the cluster domain name (**default\_realm**) is **example.com**, the value of **domain** is **hadoop.example.com**. On FusionInsight Manager, choose **Cluster > Services > KrbServer > Configurations > All Configurations** to search for and view the value of **default\_realm**.

## 5.13 Failed to Execute the Periodic Backup Task of an MRS Cluster

### Symptom

An MRS 1.7.2 cluster fails to be backed up periodically, and alarm ALM-12034 Periodic Backup Failure is displayed on the **Alarms** page of MRS Manager.

### Cause Analysis

During periodic backup of the cluster, a symbolic link is generated to connect **/srv/BigData/LocalBackup** to **/opt/Bigdata/LocalBackup**, which uses the disk space of the root directory. Because the disk space of the root directory is insufficient, backup files cannot be written into the root directory. As a result, the periodic backup of the cluster fails.

### Procedure

1. Log in to the active and standby master nodes, respectively.
2. Run the **cd /srv/BigData/** command to go to the directory where the backup file is stored.
3. Run the **unlink LocalBackup** command to delete the LocalBackup symbolic link.
4. Run the **mkdir -p LocalBackup** command to create the LocalBackup directory.
5. Run the **chown -R omm:wheel LocalBackup** command to change the user and group to which the file belongs.



6. Run the **chmod 700 LocalBackup** command to change the read and write permissions on the file.
7. Log in to MRS Manager and perform periodic backup again.

## 5.14 Failed to Download the MRS Cluster Client

### Symptom

When a user downloads the MRS cluster client from the master node to a remote node outside the MRS cluster, a message is displayed indicating that the network or parameter is incorrect.

### Cause Analysis

- The two hosts are in different VPCs.
- The password is incorrect.
- The firewall is enabled on the remote host.

### Procedure

- The two hosts are in different VPCs.  
Enable port 22 of the remote host.
- The password is incorrect.  
Check whether the password is correct. The password cannot contain special characters.
- The firewall is enabled on the remote host.  
Download the MRS cluster client to the server and run the **scp** command provided by Linux to copy the client to the remote host.

## 5.15 An Error Is Reported When a Flink Job Is Submitted in an MRS Cluster with Kerberos Authentication Enabled

### Issue

An error is reported when a user submits a Flink job in an MRS cluster with Kerberos authentication enabled.

### Symptom

Error **unable to establish the security context** is reported when a user submits a Flink sample program `./flink run /opt/client/Flink/flink/examples/streaming/WordCount.jar`.

### Cause Analysis

1. Kerberos authentication is enabled for the MRS cluster but jobs cannot be submitted. Therefore, the permission configuration is checked first. It is found

that the parameters in `/opt/client/Flink/flink/conf/flink-conf.yaml` are not correctly configured.

Figure 5-9 flink-conf.yaml configuration

```

query.proxy.query-threads: 0
security.kerberos.login.keytab: /data01/fiClient/mrs_210_new/keytab/flinkuser/user.keytab
high-availability.zookeeper.client.retry-wait: 5000
taskmanager.runtime.hashjoin-bloom-filters: false
akka.lookup.timeout: 10 s
blob.fetch.backlog: 1000
fs.obs.socket.send.buffer: 65536
blob.service.ssl.enabled: true
blob.fetch.num-concurrent: 50
query.server.query-threads: 0
akka.watch.heartbeat.interval: 10 s
security.kerberos.login.use-ticket-cache: false
fs.obs.max.total.tasks: 20
fs.obs.fast.upload: true
yarn.application-attempts: 3
akka.ask.timeout: 10 s
taskmanager.debug.memory.log-interval: 0
fs.obs.buffer.max.range: 136000
fs.obs.threads.read.core: 500
security.ssl.protocol: TLSv1.2
restart-strategy.fixed-delay.attempts: 3
akka.throughput: 15
taskmanager.network.netty.server.numThreads: 3
taskmanager.rpc.port: 32326-32390
akka.startup.timeout: 60 s
task.checkpoint.alignment.max-size: -1
taskmanager.runtime.max-fan: 128
web.ssl.enabled: false
containerized.heap-cutoff-ratio: 0.25
taskmanager.runtime.sort-spilling-threshold: 0.8
security.ssl.internal.keystore: /mrs_210_new/Flink/flink/conf//flink.keystore
fs.obs.endpoint: obs.cn-north-5.myhuaweicloud.com
    
```

keytab file path

This path is a relative path. If an absolute path is used, it needs to be configured for every node.

2. After the configuration is modified and updated, the job can be submitted again, but the "log4j:ERROR setFile(null,true) call failed" error is reported.

Figure 5-10 log4j error

```

at org.apache.flink.runtime.entrypoint.ClusterEntrypoint.<clinit>(ClusterEntrypoint.java:98)
log4j:ERROR setFile(null,true) call failed.
    
```

3. The `log4j` file shows that the user has changed the `log4j.properties` file to `log4g-cli.properties` (the name of `log4j.properties` cannot be changed randomly). As a result, an error is reported.

Figure 5-11 Viewing the log4j file

```

[emr@kwtetmr00005 flink]$ cd conf/
[emr@kwtetmr00005 conf]$ ll
total 52
-rwxr-xr-x 1 emr emr 727 Apr 15 10:26 ca.cer
-rwxr-xr-x 1 emr emr 2075 Apr 15 10:26 ca.keystore
-rwxr-xr-x 1 emr emr 358 Apr 15 08:43 client.properties
-rwxr-xr-x 1 emr emr 534 Apr 15 08:43 core-site.xml
-rwxr-xr-x 1 emr emr 773 Apr 15 10:26 flink.cer
-rwxr-xr-x 1 emr emr 5564 Apr 20 10:12 flink-conf.yaml
-rwxr-xr-x 1 emr emr 965 Apr 15 10:26 flink.csr
-rwxr-xr-x 1 emr emr 1739 Apr 15 08:43 hdfs-site.xml
-rwxr-xr-x 1 emr emr 1113 Apr 15 10:47 log4j-cli.properties
-rwxr-xr-x 1 emr emr 976 Apr 15 08:43 log4j.properties
-rwxr-xr-x 1 emr emr 709 Apr 15 08:43 log4j-yarn-session.properties
drwxrwxr-x 2 emr emr 50 Apr 17 09:50 ssl
-rwxr-xr-x 1 emr emr 1953 Apr 15 08:43 yarn-site.xml
    
```

4. After the modification, the job can be submitted properly.

Figure 5-12 Job submission

```
[ommnode-master1FtoV flink]$ flink run -m yarn-cluster /opt/client/Flink/flink/examples/streaming/WordCount.jar
2020-04-17 10:23:13.988 | WARN | [main] | Error while trying to split key and value in configuration file /opt/client/Flink/flink/conf/flink-conf.yaml:60: "security.ssl.internal.keysto
che.flink.configuration.GlobalConfiguration (GlobalConfiguration.java:179)
2020-04-17 10:23:16.844 | WARN | [main] | The short-circuit local reads feature cannot be used because libhadoop cannot be loaded. | org.apache.hadoop.hdfs.shortcircuit.DomainSocketFac
ketFactory.java:116)
2020-04-17 10:23:16.844 | WARN | [main] | The short-circuit local reads feature cannot be used because libhadoop cannot be loaded. | org.apache.hadoop.hdfs.shortcircuit.DomainSocketFac
ketFactory.java:116)
Starting execution of program
Executing WordCount example with default input data set.
Use --input to specify file input.
Printing result to stdout. Use --output to specify output path.
2020-04-17 10:23:28.548 | WARN | [main] | Error while trying to split key and value in configuration file /opt/client/Flink/flink/conf/flink-conf.yaml:60: "security.ssl.internal.keysto
che.flink.configuration.GlobalConfiguration (GlobalConfiguration.java:179)
Program execution finished
Job with JobID: job_09_cad4df48a260aaa25ba9272ce55d356 has finished
```

## Procedure

- Step 1** Determine the place to submit jobs using the client, outside the cluster or in the cluster.
1. If you want to submit jobs using the client in the cluster, switch to user **omm** before submitting jobs.
  2. If you want to submit jobs using the client outside the cluster, switch to user **root** before submitting jobs.
- Step 2** Check whether parameters in **flink-conf.yaml** are correctly set.
- Step 3** For clusters with Kerberos authentication enabled, the configuration items include keytab and principal of Kerberos.
- Download the user keytab file from the KDC server, and place the keytab file to a folder on the host of the Flink client (for example, **/home/flinkuser/keytab**).
  - Configure the following parameters in the **`\${FLINK\_HOME}/conf/flink-conf.yaml** file:
    - a. Keytab path (note that there is a space before the parameter):  
security.kerberos.login.keytab: /home/flinkuser/keytab/uer.keytab
    - b. Principal name (developer username):  
security.kerberos.login.principal:flinkuser
- Step 4** Submit the job **./flink run /opt//client/Flink/flink/examples/streaming/WordCount.jar** again and check whether the job can be submitted.
- If the job can be submitted, the permission authentication is correct. In this case, you can check other errors. In this example, the name of **log4j.properties** is changed. After the name is restored, the job can be submitted normally.
  - If the job fails to be submitted, submit a service ticket.

----End

## Related Information

For details about how to use Flink, see [Using Flink from Scratch](#).

## 5.16 An Error Is Reported When the Insert Command Is Executed on the Hive Beeline CLI

### Symptom

When the **insert into** statement is executed in Beeline of MRS Hive, the following error is reported:

```
Mapping run in Tez on Hive transactional table fails when data volume is high with error:  
"org.apache.hadoop.hive.ql.lockmgr.LockException Reason: Transaction... already aborted, Hive SQL state  
[42000]."
```

### Cause Analysis

This problem is caused by improper cluster configuration and Tez resource setting.

### Procedure

This problem can be solved by setting configuration parameters on Beeline.

**Step 1** Set the following properties to optimize performance (you are advised to change them at the cluster level):

- Set **hive.auto.convert.sortmerge.join** to **true**.
- Set **hive.optimize.bucketmapjoin** to **true**.
- Set **hive.optimize.bucketmapjoin.sortedmerge** to **true**.

**Step 2** Modify the following content to adjust the resources of Tez:

- Set **hive.tez.container.size** to the size of the Yarn container.
- Set **hive.tez.container.size** to the Yarn container size **yarn.scheduler.minimum-allocation-mb** or a smaller value (for example, a half or quarter of the Yarn container size). Ensure that the value does not exceed the **yarn.scheduler.maximum-allocation-mb** value.

----End

## 5.17 Upgrading the OS to Fix Vulnerabilities for an MRS Cluster Node

### Issue

This section describes how to upgrade the OS to fix vulnerabilities for an MRS cluster if EulerOS has vulnerabilities at the underlying layer.

### Symptom

When security software is used to test the cluster, vulnerabilities are found at the underlying layer of EulerOS.

## Cause Analysis

Services in the MRS cluster are deployed in EulerOS. Therefore, an OS upgrade is required to fix vulnerabilities.

## Procedure

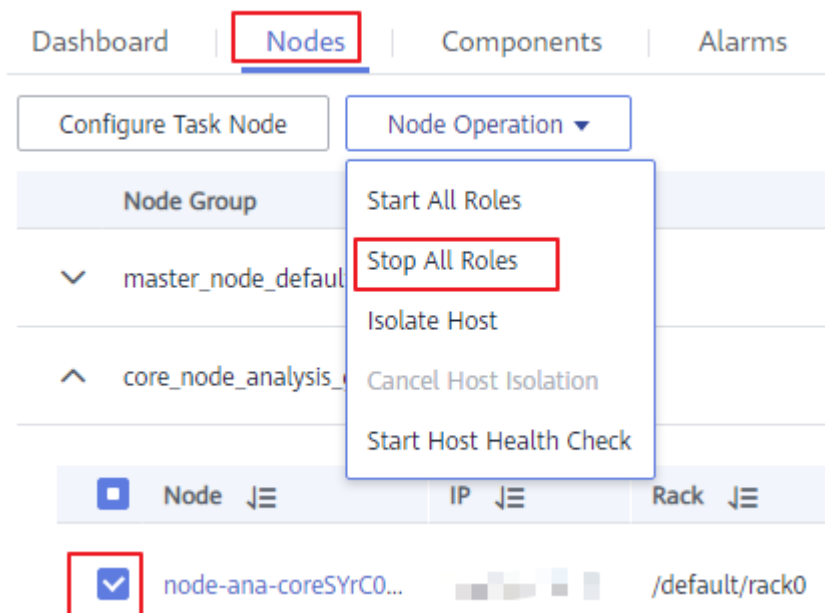
### NOTE

Before fixing the vulnerability, check whether Host Security Service (HSS) is enabled. If yes, disable HSS from monitoring the MRS cluster. After the vulnerability is fixed, enable HSS again.

**Step 1** Log in to the MRS console.

**Step 2** Click the cluster name. On the cluster details page, click the **Nodes** tab.

**Step 3** In the core node group, select a core node, click **Node Operation**, and select **Stop All Roles**.



**Step 4** Log in to the core node and configure the yum repository.

**Step 5** Run the **uname -r** or **rpm -qa |grep kernel** command to query and record the kernel version of the current node.

**Step 6** Run the **yum update -y --skip-broken --setopt=protected\_multilib=false** command to update the patch.

**Step 7** After the update is complete, query the kernel version and run the **rpm -e Old kernel version** command to delete the old kernel version.

**Step 8** On the cluster details page, click the **Nodes** tab.

**Step 9** In the core node group, click the name of the core node whose patch has been updated. The ECS console is displayed.

**Step 10** In the upper right corner of the page, click **Restart** to restart the core node.

- Step 11** On the **Nodes** tab of the cluster details page, select the core node, click **Node Operation**, and select **Start All Roles**.
- Step 12** Repeat **Step 1** to **Step 11** to upgrade other core nodes.
- Step 13** After all core nodes are upgraded, upgrade the standby master node and then the active master node. For details, see **Step 1** to **Step 11**.
- End

## 5.18 Failed to Migrate Data to MRS HDFS Using CDM

### Issue

A user failed to use CDM to migrate data from an old cluster to HDFS of a new cluster.

### Symptom

When CDM is used to import data from the source HDFS to the destination HDFS, the destination MRS cluster is faulty and the NameNode cannot be started.

The logs show that the **Java heap space** error is reported during the startup. The JVM parameter of the NameNode needs to be modified.

Figure 5-13 Fault logs

```
2020-08-27 11:44:18.327 | INFO | main | 0.02999999932947746% max memory 485.4 MB = 149.4 KB | LightweightSet.java:397
2020-08-27 11:44:18.328 | INFO | main | capacity = 214 = 16384 entries | LightweightSet.java:402
2020-08-27 11:44:18.330 | INFO | main | Using INode attribute provider: .adapter_hdfs.plugin.HWInodeAttributeProvider | FSNamesystem.java:914
2020-08-27 11:44:18.337 | INFO | main | Lock on /srv/bigdata/namenode/in_use.lock acquired by nodename 655@node-master2j6r2 | Storage.java:906
2020-08-27 11:44:18.937 | INFO | main | Planning to load image: FSImageFile{files/srv/bigdata/namenode/current/fsimage_000000000010002506, ck1txid=000000000010002506} | FSImage.java:808
2020-08-27 11:44:19.173 | INFO | main | Enable the erasure coding policy RS-6-3-1024k | ErasureCodingPolicyManager.java:410
2020-08-27 11:44:19.175 | INFO | pool-12-thread-1 | Loading 1048576 INodes. | FSImageFormatPBINode.java:336
2020-08-27 11:44:19.175 | INFO | pool-12-thread-2 | Loading 946367 INodes. | FSImageFormatPBINode.java:336
2020-08-27 11:45:33.594 | WARN | qt196612444-31-acceptor-062fa7d99-ServerConnector@20b2475a(HTTP/1.1,[http/1.1])[node-master2j6r2:9070] | | AbstractConnector.java:544
java.lang.OutOfMemoryError: Java heap space
2020-08-27 11:45:33.601 | INFO | main | Loaded FSImage in 74 seconds. | FSImageFormatProtobuf.java:295
2020-08-27 11:45:33.601 | INFO | main | Loaded image for txid 10002506 from /srv/bigdata/namenode/current/fsimage_000000000010002506 | FSImage.java:985
2020-08-27 11:45:36.045 | INFO | main | Reading org.apache.hadoop.hdfs.server.namenode.RedundantEditLogInputStream@3a94964 expecting start txid #10002507 | FSImage.java:920
```

### Cause Analysis

When the user uses CDM to migrate data, the HDFS data volume is too large. As a result, a stack exception occurs when metadata is merged.

### Procedure

- Step 1** Go to the HDFS service configuration page.
- For versions earlier than MRS 3.x: Click the cluster name on the MRS console, 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 3.x or later: Log in to FusionInsight Manager and choose **Cluster**. Click the name of the target cluster and choose **Services** > **HDFS** > **Configurations** > **All Configurations**.

**Step 2** Search for the **GC\_OPTS** parameter in **HDFS > NameNode** and increase the values of **-Xms512M** and **-Xmx512M** based on service requirements.

**Step 3** Save the configuration and restart the affected services or instances.

----End

## 5.19 Alarms Indicating Heartbeat Interruptions Between Nodes Are Frequently Generated in the MRS Cluster

### Symptom

The MRS cluster frequently reports alarms indicating that the heartbeats between active and standby Manager nodes or between active and standby DBService nodes are interrupted, or a node is faulty. As a result, Hive is occasionally unavailable, affecting upper-layer services.

### Cause Analysis

1. When the alarm is generated, the VM is restarted. The alarm is generated because the VM is restarted.

```
Macros [omm@node-master1yqIY nodeagent]$ last
omm pts/0 100.125.0.70 Thu Sep 24 10:33 still logged in
omm pts/1 100.125.0.70 Thu Sep 24 09:26 - 09:47 (00:20)
omm pts/0 100.125.0.70 Thu Sep 24 09:22 - 10:21 (00:59)
omm pts/1 100.125.0.70 Wed Sep 23 17:32 - 17:37 (00:05)
root pts/0 10.203.216.102 Wed Sep 23 17:13 - 18:35 (01:21)
omm pts/0 100.125.0.70 Wed Sep 23 16:55 - 16:56 (00:00)
omm pts/0 100.125.0.70 Wed Sep 23 16:20 - 16:25 (00:05)
reboot system boot 4.19.36-vhulk190 Wed Sep 23 16:06 still running
root pts/1 10.203.216.102 Tue Sep 22 19:13 - 19:48 (00:34)
omm pts/0 100.125.0.70 Tue Sep 22 19:08 - 20:03 (00:54)
root pts/0 10.203.216.102 Tue Sep 22 17:03 - 17:52 (00:48)
omm pts/1 100.125.0.70 Tue Sep 22 15:55 - 16:00 (00:05)
```

```
[omm@node-master2WbYp ~]$ last
omm pts/0 10.80.0.56 Thu Sep 24 11:00 still logged in
omm pts/0 10.80.0.56 Thu Sep 24 09:24 - 10:21 (00:56)
omm pts/0 10.80.0.56 Wed Sep 23 17:32 - 17:37 (00:05)
omm pts/0 10.80.0.56 Tue Sep 22 19:15 - 19:15 (00:00)
omm pts/0 10.80.0.56 Tue Sep 22 15:57 - 16:21 (00:23)
omm pts/0 10.80.0.56 Tue Sep 22 15:23 - 15:35 (00:12)
omm pts/0 10.80.0.56 Tue Sep 22 15:07 - 15:12 (00:05)
omm pts/0 10.80.0.56 Tue Sep 22 14:21 - 14:26 (00:05)
omm pts/0 10.80.0.56 Mon Sep 21 10:57 - 11:06 (00:09)
omm pts/0 10.80.0.56 Mon Sep 21 10:42 - 10:56 (00:14)
omm pts/0 10.80.0.56 Thu Sep 17 16:05 - 16:15 (00:10)
omm pts/0 10.80.0.56 Wed Sep 16 20:52 - 20:58 (00:06)
reboot system boot 4.19.36-vhulk190 Wed Sep 16 18:05 still running
omm pts/0 10.80.0.56 Wed Sep 16 15:43 - 16:10 (00:26)
omm pts/0 10.80.0.56 Wed Sep 16 14:35 - 14:53 (00:17)
omm pts/0 10.80.0.56 Wed Sep 16 14:33 - 14:33 (00:00)
omm pts/0 10.80.0.56 Wed Sep 16 14:11 - 14:29 (00:17)
omm pts/0 10.80.0.56 Wed Sep 16 14:02 - 14:09 (00:06)
omm pts/0 10.80.0.56 Wed Sep 16 11:56 - 12:04 (00:08)
omm pts/0 10.80.0.56 Wed Sep 16 11:26 - 11:31 (00:04)
omm pts/0 10.80.0.56 Wed Sep 16 11:09 - 11:24 (00:15)
root pts/0 10.203.230.193 Mon Sep 14 15:54 - 16:30 (00:35)
root pts/0 10.203.172.29 Fri Sep 11 17:15 - 17:45 (00:30)
root pts/0 10.203.172.29 Fri Sep 11 16:53 - 17:12 (00:19)
root tty1 Fri Sep 11 16:23 - 17:25 (01:01)
reboot system boot 4.19.36-vhulk190 Fri Sep 11 10:07 still running
reboot system boot 4.19.36-vhulk190 Thu Aug 27 16:41 still running
root tty1 Thu Aug 20 09:46 - 10:17 (00:30)
reboot system boot 4.19.36-vhulk190 Wed Aug 19 17:48 still running
reboot system boot 4.19.36-vhulk190 Wed Aug 19 17:46 still running
```

2. According to the OS analysis, the cause of the VM restart is that the node does not have available memory. Memory overflow triggers oom-killer. When the process is invoked, the process enters the **disk sleep** state. As a result, the VM restarts.

```
mem info:
[344766.903734] MemTotal: 32397404 kB ← Total memory
MemFree: 160404 kB
MemAvailable: 31668 kB
Buffers: 2172 kB
Cached: 2768904 kB
SwapCached: 0 kB
Active: 30328872 kB ← Used by the user
Inactive: 1035844 kB
Active(anon): 30320852 kB
Inactive(anon): 1004376 kB
Active(file): 8020 kB
Inactive(file): 31468 kB
Unevictable: 0 kB
Mlocked: 0 kB
[344766.903738] SwapTotal: 0 kB
SwapFree: 0 kB
```



```

[344766.904470] 20444      1 212884K      104K      S (sleeping) /sbin/agetty -o -p -- \u --noclear tty1 linux
[344766.904474] 15011  9241  845712K      1948K      S (sleeping) gaussdb: wal sender process REPLICATION node-masterlyqiy(30753) s
[344766.904477] 20394  9241  866276K      326202K    D (disk sleep) gaussdb: OMM OMM localhost(35218) FARSE
[344766.904480] 20399  9241  867524K      326732K    D (disk sleep) gaussdb: OMM OMM localhost(35222) FARSE
[344766.904484] 29394      1 253256K      1852K      S (sleeping) /usr/sbin/sssd -D
[344766.904487] 29453 29394 253144K      2620K      R (running) /usr/libexec/sss/sss_be --domain implicit_files --uid 0 --gid 0 --logger=journald
[344766.904491] 29454 29384 258292K      4004K      S (sleeping) /usr/libexec/sss/sss_be --domain default --uid 0 --gid 0 --logger=journald
[344766.904494] 29512 29384 283272K      2112K      S (sleeping) /usr/libexec/sss/sss_nss --uid 0 --gid 0 --logger=journald
[344766.904498] 29513 29384 243880K      1680K      D (disk sleep) /usr/libexec/sss/sss_pam --uid 0 --gid 0 --logger=journald
[344766.904501] 29527      1 5500276K     323624K    S (sleeping) /opt/Bigdata/jdk1.8.0_212/bin/java -cp
/opt/Bigdata/MRS_2.1.0/1_21_JDBCServer/etc/1/opt/Bigdata/security:/opt/Bigdata/MRS_2.1.0/install/FusionInsight-Spark-2.3.2/spark/sbin/./jars/* -Dlog4
-Djava.security.auth.Login.config=/o
[344766.904505] 7855  9241  846668K      23736K      S (sleeping) gaussdb: OMM OMM localhost(46200) idle
[344766.904509] 25941  9241  859332K      323464K    D (disk sleep) gaussdb: OMM OMM localhost(48556) idle
[344766.904512] 25951  9241  857892K      319088K    D (disk sleep) gaussdb: OMM OMM localhost(48558) FARSE
[344766.904516] 26004  9241  867192K      324348K    D (disk sleep) gaussdb: OMM OMM localhost(48562) idle
[344766.904519] 26108  9241  857940K      323228K    D (disk sleep) gaussdb: OMM OMM localhost(48564) FARSE
[344766.904523] 26156  9241  858120K      324052K    D (disk sleep) gaussdb: OMM OMM localhost(48570) FARSE
[344766.904527] 26165  9241  846212K      322884K    D (disk sleep) gaussdb: OMM OMM localhost(48576) FARSE
[344766.904531] 26172  9241  858180K      322896K    D (disk sleep) gaussdb: OMM OMM localhost(48578) FARSE
[344766.904534] 26212  9241  857932K      323148K    D (disk sleep) gaussdb: OMM OMM localhost(48580) FARSE
[344766.904538] 26309  9241  859160K      321728K    D (disk sleep) gaussdb: OMM OMM localhost(48582) FARSE
[344766.904541] 26362  9241  866236K      322212K    D (disk sleep) gaussdb: OMM OMM localhost(48584) FARSE
[344766.904545] 26399  9241  866408K      323184K    D (disk sleep) gaussdb: OMM OMM localhost(48588) FARSE
[344766.904548] 26399  9241  857844K      321616K    D (disk sleep) gaussdb: OMM OMM localhost(48592) FARSE
[344766.904551] 26404  9241  859044K      322592K    D (disk sleep) gaussdb: OMM OMM localhost(48596) FARSE
[344766.904555] 26415  9241  857756K      322528K    D (disk sleep) gaussdb: OMM OMM localhost(48600) FARSE
[344766.904558] 26450  9241  858768K      323668K    D (disk sleep) gaussdb: OMM OMM localhost(48606) FARSE
[344766.904562] 26482  9241  858072K      323340K    D (disk sleep) gaussdb: OMM OMM localhost(48608) FARSE
[344766.904565] 26608  9241  858024K      323504K    D (disk sleep) gaussdb: OMM OMM localhost(48610) FARSE
[344766.904568] 27449  9241  846276K      323472K    D (disk sleep) gaussdb: OMM OMM localhost(48632) FARSE
[344766.904573] 30030      1 387064K      17424K      R (running) /opt/Bigdata/MRS_2.1.0/install/FusionInsight-Hue-3.11.0/hue/build/env/bin/python2.7
/opt/Bigdata/MRS_2.1.0/install/FusionInsight-Hue-3.11.0/hue/build/env/bin/supervisor -p /opt/Bigdata/MRS_2.1.0/install/FusionInsight-Hue-3.11.0/hue/cnf/
[344766.904726] 874  4953  1484K      8K      D (disk sleep) /bin/sh /opt/Bigdata/nodeagent/bin/scriptlauncher.sh /opt/Bigdata/MRS_2.1.0/install/dsbservice/sh
[344766.904729] 875 26044  1488K      12K      D (disk sleep) /bin/sh /opt/Bigdata/nodeagent/bin/scriptlauncher.sh
/opt/Bigdata/MRS_2.1.0/install/FusionInsight-Hadoop-3.11/hadoop/sbin/yarn-resourcemanager-check.sh
[344766.904732] 876 10755 752240K  670728K    D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Dprocess.name=nodeagent
-Dbeetle.application.home.path=/opt/Bigdata/security/config -Dsun.rmi.transport.tcp.responseTimeout=60000 -Djava.library.path=/opt/Bigdata/nodeagent/lib
-XX:ErrorFile=/var/log/Bigdata/nodeagent
[344766.904735] 878 17629 8616200K 1124612K D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Djava.security.egd=file:/dev/./urandom -Dprocess.name=contn
-Datack.conf.dir=/Doocontroller.home=/opt/Bigdata/om-0.0.1 -Dbeetle.application.home.path=/opt/Bigdata/om-0.0.1/etc/om -Dorg.terracotta.quartz.skipUpdate
[344766.904738] 879 7057  1484K      8K      D (disk sleep) /bin/sh /opt/Bigdata/nodeagent/bin/scriptlauncher.sh
/opt/Bigdata/MRS_2.1.0/install/FusionInsight-Flume-1.6.0/flume/bin/flume-check-service.sh
[344766.904741] 880 2535  1488K      12K      D (disk sleep) /bin/sh /opt/Bigdata/nodeagent/bin/scriptlauncher.sh /usr/bin/head -1 /opt/Bigdata/tmp/hadoop-
[344766.904744] 881 9760 752240K  670728K    D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Dprocess.name=nodeagent
-Dbeetle.application.home.path=/opt/Bigdata/security/config -Dsun.rmi.transport.tcp.responseTimeout=60000 -Djava.library.path=/opt/Bigdata/nodeagent/lib
-XX:ErrorFile=/var/log/Bigdata/nodeagent
[344766.904746] 882 3895 752240K  670728K    D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Dprocess.name=nodeagent
-Dbeetle.application.home.path=/opt/Bigdata/security/config -Dsun.rmi.transport.tcp.responseTimeout=60000 -Djava.library.path=/opt/Bigdata/nodeagent/lib
-XX:ErrorFile=/var/log/Bigdata/nodeagent
[344766.904748] 883 3665 752240K  670728K    D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Dprocess.name=nodeagent
-Dbeetle.application.home.path=/opt/Bigdata/security/config -Dsun.rmi.transport.tcp.responseTimeout=60000 -Djava.library.path=/opt/Bigdata/nodeagent/lib
-XX:ErrorFile=/var/log/Bigdata/nodeagent
[344766.904751] 885 843 752240K  670728K    D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Dprocess.name=nodeagent
-Dbeetle.application.home.path=/opt/Bigdata/security/config -Dsun.rmi.transport.tcp.responseTimeout=60000 -Djava.library.path=/opt/Bigdata/nodeagent/lib
-XX:ErrorFile=/var/log/Bigdata/nodeagent
[344766.904753] 886 5536 752240K  670728K    D (disk sleep) /opt/Bigdata/jdk1.8.0_212/bin/java -Dprocess.name=nodeagent
-Dbeetle.application.home.path=/opt/Bigdata/security/config -Dsun.rmi.transport.tcp.responseTimeout=60000 -Djava.library.path=/opt/Bigdata/nodeagent/lib
-XX:ErrorFile=/var/log/Bigdata/nodeagent
[344766.904754] Mem-Info:
[344766.904757] active anon:7580213 inactive anon:251094 isolated anon:0

```

3. Check the processes that occupy the memory. It is found that the processes that occupy the memory are normal service processes.

Conclusion: The VM memory cannot meet service requirements.

## Procedure

- You are advised to expand the node memory.
- You are advised to disable unnecessary services.

## 5.20 High Memory Usage of the PMS Process

### Issue

What can I do if the memory usage of the active Master node is high?

### Symptom

The memory usage of the active Master node is high. The **top -c** command output shows that the following idle processes occupy a large amount of memory:

12180	ommdba	20	0	1395492	1.180g	1.082g	S	0.0	3.8	23:14.29	gaussdb:	OMM	OMM	localhost(60598)	idle
14828	ommdba	20	0	1395904	1.180g	1.081g	S	0.0	3.8	23:17.08	gaussdb:	OMM	OMM	localhost(60698)	idle
15016	ommdba	20	0	1395840	1.180g	1.081g	S	0.0	3.8	23:11.19	gaussdb:	OMM	OMM	localhost(60824)	idle
14943	ommdba	20	0	1395900	1.180g	1.081g	S	0.0	3.8	23:14.76	gaussdb:	OMM	OMM	localhost(60764)	idle
14908	ommdba	20	0	1395840	1.180g	1.081g	S	0.0	3.8	23:15.18	gaussdb:	OMM	OMM	localhost(60738)	idle
14953	ommdba	20	0	1395824	1.180g	1.081g	S	0.0	3.8	23:15.96	gaussdb:	OMM	OMM	localhost(60770)	idle
14995	ommdba	20	0	1395560	1.180g	1.081g	S	0.0	3.8	23:13.28	gaussdb:	OMM	OMM	localhost(60812)	idle
15062	ommdba	20	0	1395820	1.180g	1.081g	S	0.0	3.8	23:16.12	gaussdb:	OMM	OMM	localhost(60868)	idle
15064	ommdba	20	0	1395512	1.180g	1.081g	S	0.0	3.8	23:13.33	gaussdb:	OMM	OMM	localhost(60870)	idle
14973	ommdba	20	0	1395528	1.180g	1.081g	S	0.0	3.8	23:12.74	gaussdb:	OMM	OMM	localhost(60790)	idle
14835	ommdba	20	0	1395536	1.180g	1.081g	S	0.0	3.8	23:17.39	gaussdb:	OMM	OMM	localhost(60704)	idle
14822	ommdba	20	0	1395524	1.180g	1.081g	S	0.0	3.8	23:13.80	gaussdb:	OMM	OMM	localhost(60692)	idle
14991	ommdba	20	0	1395808	1.180g	1.081g	S	0.0	3.8	23:17.96	gaussdb:	OMM	OMM	localhost(60808)	idle
14975	ommdba	20	0	1395812	1.180g	1.081g	S	0.0	3.8	23:12.57	gaussdb:	OMM	OMM	localhost(60792)	idle
15038	ommdba	20	0	1395520	1.180g	1.081g	S	0.0	3.8	23:12.75	gaussdb:	OMM	OMM	localhost(60846)	idle
14919	ommdba	20	0	1395540	1.180g	1.081g	S	0.0	3.8	23:11.58	gaussdb:	OMM	OMM	localhost(60744)	idle
14832	ommdba	20	0	1395476	1.180g	1.081g	S	0.0	3.8	23:13.11	gaussdb:	OMM	OMM	localhost(60702)	idle
14989	ommdba	20	0	1395500	1.180g	1.081g	S	0.0	3.8	23:15.63	gaussdb:	OMM	OMM	localhost(60806)	idle
14979	ommdba	20	0	1395448	1.180g	1.081g	S	0.0	3.8	23:13.17	gaussdb:	OMM	OMM	localhost(60796)	idle
15047	ommdba	20	0	1395512	1.180g	1.081g	S	0.0	3.8	23:12.10	gaussdb:	OMM	OMM	localhost(60854)	idle
14977	ommdba	20	0	1395496	1.180g	1.081g	S	0.0	3.8	23:16.90	gaussdb:	OMM	OMM	localhost(60794)	idle
15028	ommdba	20	0	1395800	1.180g	1.081g	S	0.0	3.8	23:09.35	gaussdb:	OMM	OMM	localhost(60836)	idle

## Cause Analysis

- PostgreSQL cache: In addition to common execution plan cache and data cache, PostgreSQL provides cache mechanisms such as **catalog** and **relation** to improve the efficiency of generating execution plans. In the persistent connection scenario, some of the caches are not released. As a result, the persistent connection may occupy a large amount of memory.
- PMS is a monitoring process of MRS. This process frequently creates table partitions or new tables. The PostgreSQL caches the metadata of the objects accessed by the current session, and the connections in the database connection pool of the PMS exist for a long time. Therefore, the memory occupied by the connections gradually increases.

## Procedure

**Step 1** Log in to the active Master node as user **root**.

**Step 2** Run the following command to query the PMS process ID:

```
ps -ef | grep =pmsd |grep -v grep
```

**Step 3** Run the following command to stop the PMS process. In the command, **PID** indicates the PMS process ID obtained in [Step 2](#).

```
kill -9 PID
```

**Step 4** Wait for the PMS process to automatically start.

It takes 2 to 3 minutes to start PMS. PMS is a monitoring process. Restarting PMS does not affect big data services.

----End

## 5.21 High Memory Usage of the Knox Process

### Issue

The memory usage of the knox process is high.

## Symptom

The memory usage of the active master node is high. The **top -c** command output shows that the memory usage of the **knox** process exceeds 4 GB.

## Cause Analysis

The memory is not separately configured for the **knox** process. The process automatically allocates available memory based on the system memory size. As a result, the **knox** process occupies a large amount of memory.

## Procedure

- Step 1** Log in to the master nodes as user **root**.
- Step 2** Open the **/opt/knox/bin/gateway.sh** file. Search for **APP\_MEM\_OPTS**, and set its value to **-Xms3072m -Xmx4096m**.
- Step 3** Log in to Manager, find the IP address of the active master node (that is, the node with a solid star before the hostname) in the host list, and then log in to the node.
- Step 4** Run the following commands to restart the process:

```
su - omm
```

```
sh /opt/knox/bin/restart-knox.sh
```

```
----End
```

## 5.22 It Takes a Long Time to Access HBase from a Client Outside a Security Cluster

### Symptom

A user creates an MRS security cluster, installs a cluster client on a node outside the cluster, and runs the **hbase shell** command to access HBase. It is found that the access to HBase is very slow.

### Cause Analysis

Kerberos authentication is required for a security cluster. You need to configure the **hosts** file on the client node to ensure that the access speed is not affected.

### Procedure

Copy the content of the **hosts** file on the cluster node to the **hosts** file on the node where the client is installed.

## 5.23 Failed to Submit Jobs

### Symptom

Jobs cannot be submitted on the DataArts Studio or MRS console.

### Impact

Jobs cannot be submitted, and services are interrupted.

### Procedure

**Step 1** Locate the cause of the exception.

View the error code in the job log and check whether the error is reported for the APIG or MRS.

- For an error of the public APIG (starting with **APIGW**), contact public APIG maintenance personnel.
- For an MRS error, go to the next step.

**Step 2** Check the running status of services and processes.

1. Log in to Manager and check whether a service fault occurs. If a job-related service fault or an underlying service fault occurs, rectify the fault.
2. Check whether a critical alarm is generated.
3. Log in to the active master node.
4. Run the following command to check whether the OMS status is normal and whether the Executor and Knox processes on the active OMS node are normal. Knox is deployed on dual active nodes, and Executor is deployed on a single active node.

**/opt/Bigdata/om-0.0.1/sbin/status-oms.sh**

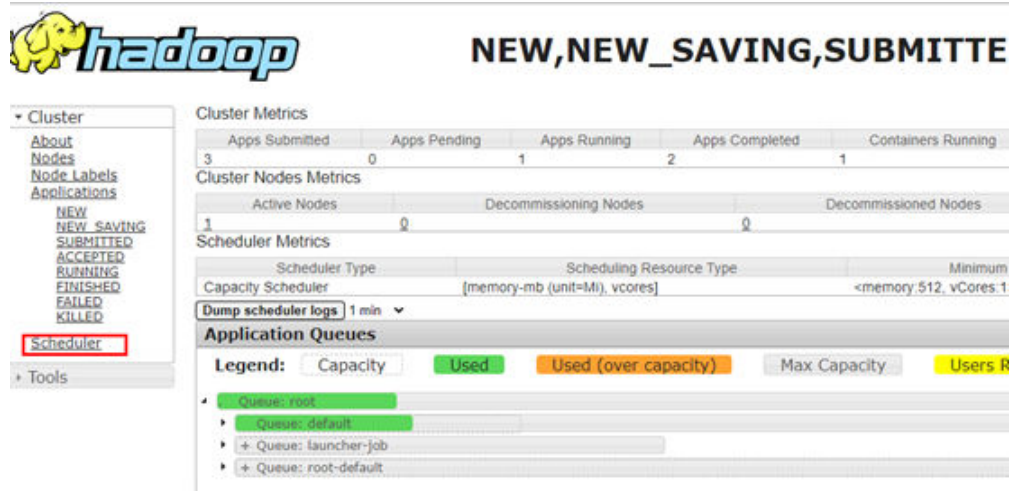
5. Run the **jmap -heap PID** command as user **omm** to check the memory usage of the Knox and Executor processes. If the old-generation memory usage is 99.9%, the memory overflow occurs.
  - Run the **netstat -anp | grep 8181 | grep LISTEN** command to query the PID of the Executor process.
  - Run the **ps -ef|grep knox | grep -v grep** command to query the PID of the Knox process.

If the memory overflows, run the **jmap -dump:format=b,file=/home/omm/temp.bin PID** command to export the memory information and restart the process.

6. View the native Yarn page to check the queue resource usage and whether the job has been submitted to Yarn.

On the native Yarn page, choose **Components > Yarn > ResourceManager WebUI > ResourceManager (Active)**.

Figure 5-14 Queue resource usage on the Yarn page



**Step 3** Locate the fault causing the job submission failure.

1. Log in to the MRS console and click the cluster name to go to the cluster details page.
2. On the **Jobs** page, locate the row that contains the target job and click **View Log** in the **Operation** column.
3. If there is no log or the log information is not detailed, copy the job ID in the **Name/ID** column.
4. Run the following command on the active OMS node to check whether the job request is sent to Knox. If the request is not sent to Knox, Knox may be faulty. In this case, restart Knox to rectify the fault.

```
grep "mrsjob" /var/log/Bigdata/knox/logs/gateway-audit.log | tail -10
```

5. Search for the job ID in the Executor log and view the error information.  
Log file path: `/var/log/Bigdata/executor/logs/exe.log`
6. Modify the `/opt/executor/webapps/executor/WEB-INF/classes/log4j.properties` file to enable the DEBUG log of the Executor. Submit the test job and view the Executor log. Check the error reported during job submission.

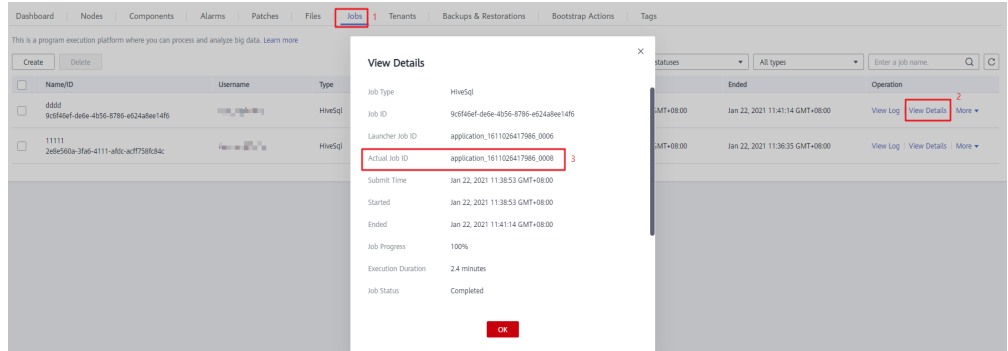
```
Log file path: /var/log/Bigdata/executor/logs/exe.log
```

7. If an error occurs in the Executor, run the following command to print the jstack information of the Executor and check the current execution status of the thread:

```
jstack PID > xxx.log
```

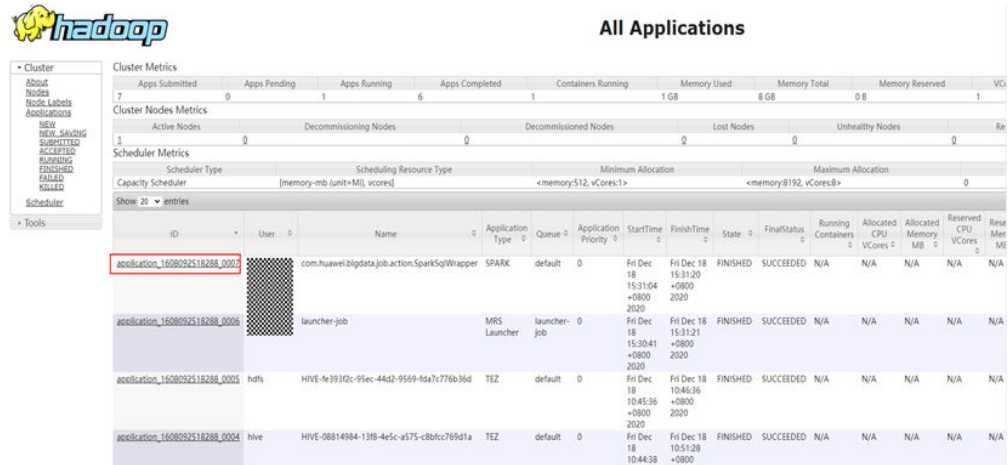
8. On the cluster details page, click the **Jobs** tab. Locate the row that contains the target job, and click **View Details** in the **Operation** column to obtain the actual job ID (application ID).

Figure 5-15 Job details



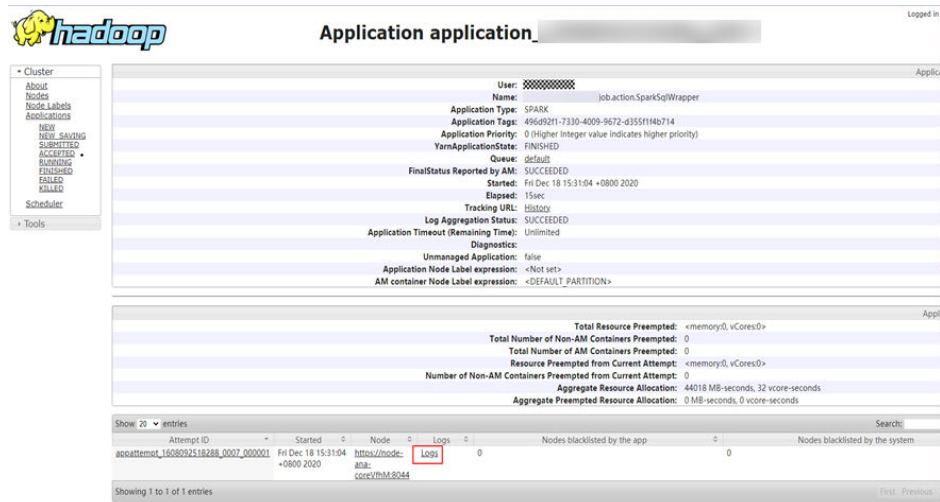
- On the cluster details page, choose **Components > Yarn > ResourceManager WebUI > ResourceManager (Active)**. On the native Yarn page that is displayed, click the application ID.

Figure 5-16 Yarn applications



- View logs on the task details page.

Figure 5-17 Task logs



----End

## 5.24 OS Disk Space Is Insufficient Due to Oversized HBase Log Files

### Issue

The space of the `/var/log` partition on the system disk is insufficient.

### Symptom

The `/var/log/Bigdata/hbase/*/hbase-omm-*.out` log file is too large, causing insufficient space of the `/var/log` partition on the system disk.

### Cause Analysis

During the long-term running of HBase, the OS periodically deletes the `/tmp/.java_pid*` files created by the JVM. The HBase memory monitoring uses the `jinfo` command, which depends on the `/tmp/.java_pid*` file. If the file does not exist, the `jinfo` command runs `kill -3` to print the jstack information to the `.out` log file. As a result, the `.out` log file becomes oversized as time goes by.

### Procedure

On each node hosting the HBase instance, deploy a scheduled task to periodically clear the `.out` log file. For example, log in to the HBase instance node and run the `crontab -e` command to add a scheduled task to clear the `.out` log file at 00:00:00 every day.

```
crontab -e
```

```
00 00 * * * for file in `ls /var/log/Bigdata/hbase/*/hbase-omm-*.out`; do echo "" > $file; done
```

#### NOTE

If large `.out` files are generated frequently, you can clear the files multiple times every day or adjust the automatic clearing policy of the OS.

## 5.25 OS Disk Space Is Insufficient Due to Oversized HDFS Log Files

### Issue

The space of the `/var/log` partition on the system disk is insufficient.

### Symptom

The `/var/log/Bigdata/hdfs/*/hdfs-omm-*.out` log file is too large, causing insufficient space of the `/var/log` partition on the system disk.

## Cause Analysis

During the long-term running of HDFS, the OS periodically deletes the `/tmp/.java_pid*` files created by the JVM. The HDFS memory monitoring uses the `jinfo` command, which depends on the `/tmp/.java_pid*` file. If the file does not exist, the `jinfo` command runs `kill -3` to print the jstack information to the `.out` log file. As a result, the `.out` log file becomes oversized as time goes by.

## Procedure

On each node hosting the HDFS instance, deploy a scheduled task to periodically clear the `.out` log file. For example, log in to the HDFS instance node and run the `crontab -e` command to add a scheduled task to clear the `.out` log file at 00:00:00 every day.

`crontab -e`

```
00 00 * * * for file in `ls /var/log/Bigdata/hdfs*/hdfs-omm-*.out`; do echo "" > $file; done
```

 NOTE

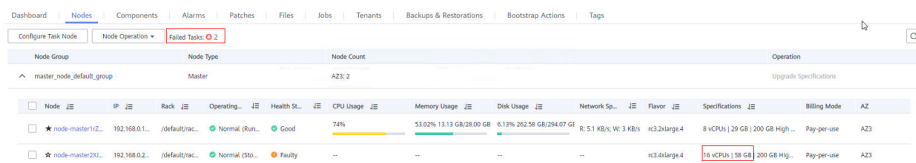
If large `.out` files are generated frequently, you can clear the files multiple times every day or adjust the automatic clearing policy of the OS.

## 5.26 An Exception Occurs During Specifications Upgrade of Nodes in an MRS Cluster

### Symptom

A component may fail to be started or encounter an exception after nodes specifications are upgraded. As a result, an exception occurs in the upgrade. The symptoms are as follows:

- The specifications of the failed node have been upgraded.



Node Group	Node Type	Node Count	Operation
master_node_default_group	Master	AZ3: 2	Upgrade Specifications
node-master1z...	Normal (Run...)	Good	74%
node-master2z...	Normal (Sta...)	Failed	...

- A new specifications upgrade failure is displayed on the **Manage Failed Tasks** page.



**Manage Failed Tasks**

The following table lists failed cluster instances.

Cluster Name	ID	Task Type	Failed	Operation
mrs_lzyeuan	3acfdueb-fd2c-49d3-83d0-6685...	Scale Up ...	Failed	Delete
mrs_lzyeuan	3acfdueb-fd2c-49d3-83d0-6685...	Scale Up ...	Failed	Delete

**(MRS-3602) Failed due to scale up master node.**  
Cluster ID: 3acfdueb-fd2c-49d3-83d0-6685df4112a3.  
java.util.concurrent.TimeoutException: Wait cluster instance good timeout.



- If IAM users have been synchronized, you can view abnormal roles on the **Components** page.

Service	Operating Status	Health Status	Configuration Status	Roles	Operation
DBService	Started	Good	Synchronized	DBServer: 2 (Unknown: 1)	Start   Stop   Restart
HBase	Started	Good	Synchronized	RegionServer: 1 ThriftServer: 2 (Unknown: 1) RESTServer: 2 (Unknown: 1) HMaster: 2 (Unknown: 1)	Start   Stop   Restart
HDFS	Started	Good	Synchronized	Zkfc: 2 (Unknown: 1) DataNode: 1 JournalNode: 3 (Unknown: 1) NameNode: 2 (Unknown: 1)	Start   Stop   Restart
Hive	Started	Good	Synchronized	MetaStore: 2 (Unknown: 1) WebHCat: 2 (Unknown: 1) HiveServer: 2 (Unknown: 1)	Start   Stop   Restart
Hue	Started	Good	Synchronized	Hue: 2 (Unknown: 1)	Start   Stop   Restart
KIBServer	Started	Good	Synchronized	KerberosServer: 2 (Unknown: 1) KerberosAdmin: 2 (Unknown: 1)	Start   Stop   Restart
LdapServer	Started	Good	Synchronized	SlapdServer: 2 (Unknown: 1)	Start   Stop   Restart
MapReduce	Started	Good	Synchronized	JobHistoryServer: 1	Start   Stop   Restart
meta	Started	Good	Synchronized	meta: 3 (Unknown: 1)	Start   Stop   Restart
Presto	Started	Good	Synchronized	Coordinator: 1 Worker: 1	Start   Stop   Restart

If they are not synchronized, you can view abnormal roles on the Manager page of this cluster.

## Cause Analysis

The failure is caused by the cluster component exception.

## Procedure

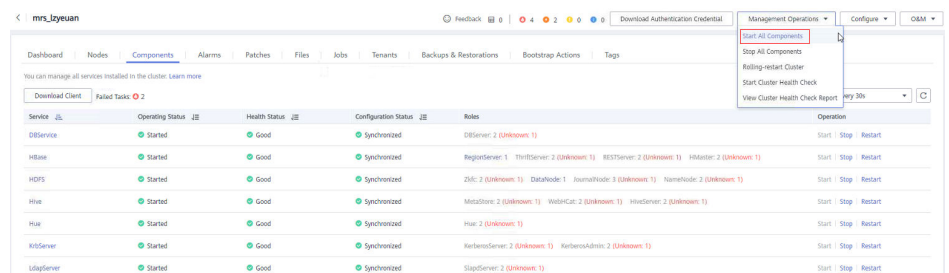
- Method 1:
  - Go to the **Components** page and check the service health status and operating status. If any service is not started, click the service to go to its details page.

- On the **Instances** page, select the instance that is not started, click **More**, and select **Start Instance**.

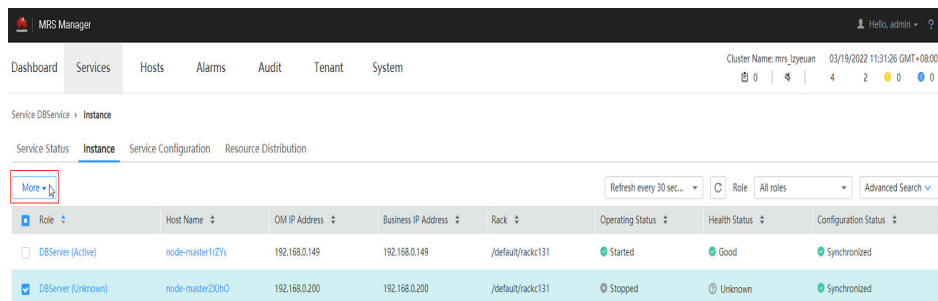
Role	Host Name	OM IP Address	Business IP Address	Rack	Operating Status	Health Status	Configuration Status
DBServer(Active)	node-master1-ZYS	192.168.0.149	192.168.0.149	/default/ rack131	Started	Good	Synchronized
DBService(Unknown)	node-master2XHO	192.168.0.200	192.168.0.200	/default/ rack131	Stopped	Unknown	Synchronized

**NOTE**

If there are many abnormal roles, you can choose **Start All Components** from the **Management Operations** drop-down list.



- **Method 2:**  
Go to the Manager page of the cluster and check whether there are instances that are not started. If yes, start the instances.



- For other exceptions that cannot be resolved, contact technical support.

## 5.27 Failed to Delete a New Tenant on FusionInsight Manager

### Symptom

A user fails to delete a tenant created on the **Tenant Resources** page of FusionInsight Manager, and an error message is displayed.

### Cause Analysis

When a tenant is created, its role is generated. The role will be deleted first when the tenant is deleted. If the component that supports permission configuration is abnormal, the resource permission of the role fails to be deleted.

### Procedure

- Step 1** Log in to FusionInsight Manager and choose **System > Permission > Role**.
- Step 2** Click **Create Role**. In the **Configure Resource Permission** area, click the cluster name to check the components available for resource permission configuration.
- Step 3** Choose **Cluster > Services** and check that the running status of these components is **Normal**.

**Step 4** (Optional) If the running status is not **Normal**, start or repair the component until its running status becomes **Normal**.

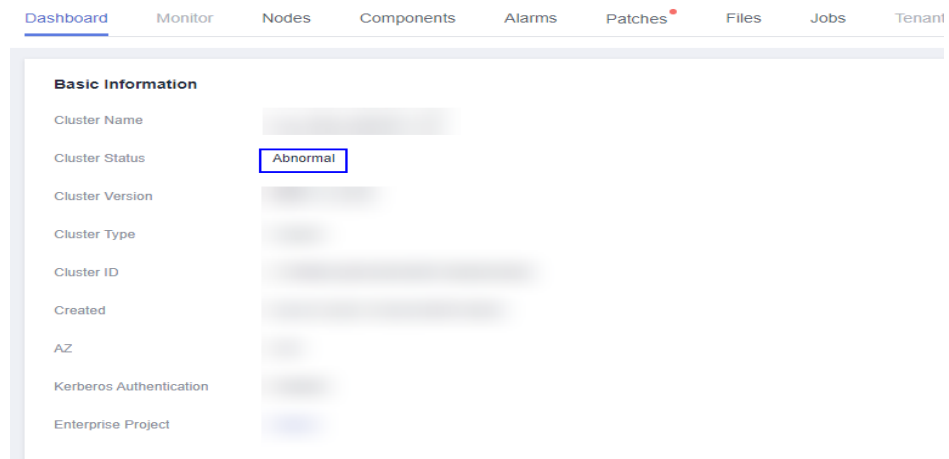
**Step 5** Delete the tenant again.

----End

## 5.28 MRS Cluster Becomes Unavailable After the VPC Is Changed

### Symptom

In an MRS cluster, after the VPCs of all nodes are changed on the ECS console, the cluster status becomes abnormal.



All services are unavailable. The Hive Beeline reports the following error.

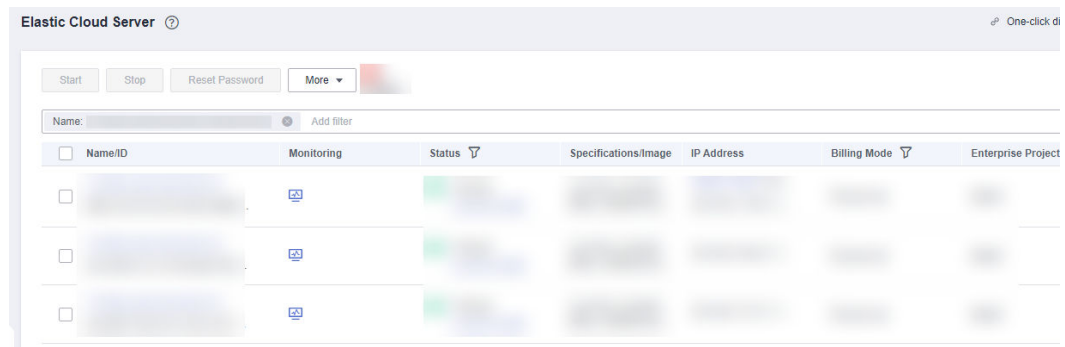
```
2021-09-15 10:54:26,741 ERROR Imps.CuratorFrameworkImpl: Background operation retry gave up
org.apache.zookeeper KeeperException$ConnectionLossException: KeeperErrorCode = ConnectionLoss
    at org.apache.zookeeper.KeeperException.create(KeeperException.java:187)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.checkBackgroundRetry(CuratorFrameworkImpl.java:862)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.performBackgroundOperation(CuratorFrameworkImpl.java:998)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.backgroundOperationsLoop(CuratorFrameworkImpl.java:943)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.access$388(CuratorFrameworkImpl.java:66)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl$4.call(CuratorFrameworkImpl.java:346)
    at java.util.concurrent.FutureTask.run(FutureTask.java:266)
    at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.access$291(ScheduledThreadPoolExecutor.java:118)
    at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run(ScheduledThreadPoolExecutor.java:293)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:748)
2021-09-15 10:54:26,764 ERROR Imps.CuratorFrameworkImpl: Background retry gave up
org.apache.curator.CuratorConnectionLossException: KeeperErrorCode = ConnectionLoss
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.performBackgroundOperation(CuratorFrameworkImpl.java:972)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.backgroundOperationsLoop(CuratorFrameworkImpl.java:943)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl.access$388(CuratorFrameworkImpl.java:66)
    at org.apache.curator.framework.imps.CuratorFrameworkImpl$4.call(CuratorFrameworkImpl.java:346)
    at java.util.concurrent.FutureTask.run(FutureTask.java:266)
    at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.access$291(ScheduledThreadPoolExecutor.java:118)
    at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run(ScheduledThreadPoolExecutor.java:293)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:748)
```

### Cause Analysis

MRS does not support VPC change. After the VPC is changed, the internal IP address of the node changes, but the configuration file and database still use the original IP address. As a result, functions such as cluster communication are abnormal, and the cluster status is also abnormal. Therefore, to restore the cluster, you need to change the VPC back and ensure that the IP address maps that in the **hosts** file.

## Procedure

- Step 1** Log in to the Master1 node and run the **ifconfig** command to view the new VPC. Run the **cat /etc/hosts** command to check the IP address recorded in the **hosts** file before the VPC change.
- Step 2** Log in to the MRS console and view the cluster ID and VPC on the **Dashboard** page of the cluster.
- Step 3** Log in to the ECS console, select **Name** in the search box, and enter the MRS cluster ID to search for all nodes in the MRS cluster.




- Step 4** In the **Operation** column of the MRS cluster node, click **More** and choose **Manage Network > Change VPC**.

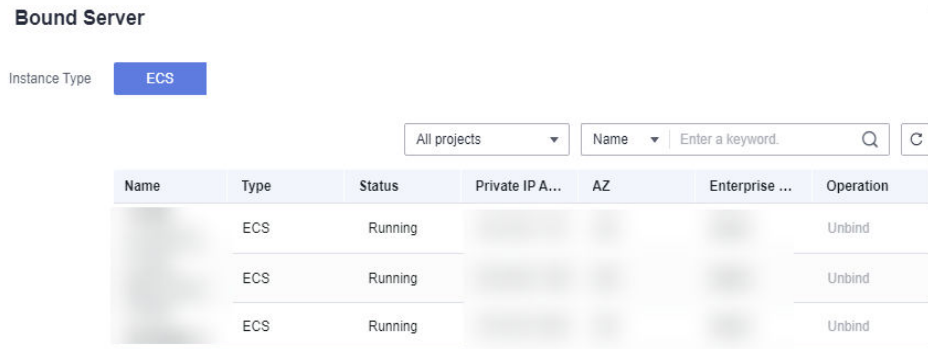
### NOTE

- You need to change the VPC for each node.
- When changing the VPC, ensure that the VPC, subnet, and security group are the same as those in the initial cluster configuration.
- Set **Private IP Address** to **Assign new** and enter the IP address of the node queried in [Step 1](#).

- Step 5** After the change is successful, click the node name, switch to the **Network Interfaces** tab, and enable **Source/Destination Check** again.

- Step 6** Perform the following steps to rebind the virtual IP address to the master node of the cluster:

1. Log in to the MRS console and access the MRS cluster. On the **Dashboard** page, click  next to **Access Manager**, set **Access Mode** to **Direct Connect**, and record the floating IP address of the cluster. View and take note of the subnet in **Default Subnet**.
2. Log in to the VPC console, choose **Virtual Private Cloud > Subnets**, and search for the subnet of the MRS cluster.
3. Click the subnet name, click the **IP Addresses** tab, and search for the floating IP address of the MRS cluster.
4. Click **Bind to Server** in the **Operation** column of the floating IP address. On the **Bind to Server** page, select the master node of the MRS cluster. After the binding is successful, the following figure is displayed.



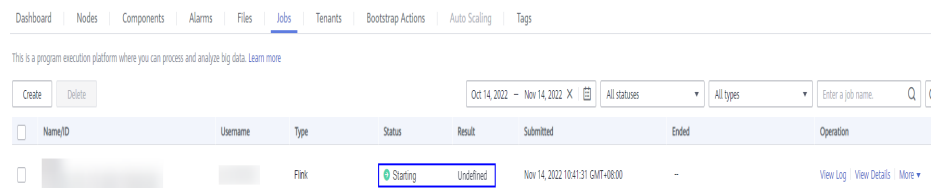
**Step 7** Wait for the cluster to restore.

----End

## 5.29 Failed to Submit Jobs on the MRS Console

### Issue

On the **Jobs** page of a cluster on the MRS console, the status of a job is **Starting** but its result is **Undefined**, as shown in the following figure. The job fails to be submitted to Yarn.



### Cause Analysis

The job management function of the console is scheduled and executed by the cluster management module Executor. Therefore, you need to check the Executor to locate the job submission failure. In normal cases, after a job is added, the job management function on the console automatically starts two jobs on Yarn. One job is submitted to the launcher-job queue, which is an auxiliary job queue. The other is submitted to the queue where the job is actually executed, for example, the default queue.

Check the Executor log (`/var/log/executor/exe.log` on the active master node). It is found that the keytab authentication file fails to be downloaded because the password of the user who submits the Flink job is changed or expires. As a result, the job is not submitted to the launcher-job queue.

### Procedure

**Step 1** Reset the password of the user who submits the job.

Log in to Manager, choose **System** > **Permission** > **User**. In the **Operation** column of the IAM user who submits the job, click **More**, select **Initialize Password**, and perform operations as prompted. After the initialization is complete, you need to log in to Manager as this user.

- Step 2** Log in to the MRS console and access the MRS cluster. On the **Dashboard** page, click **Synchronize** on the right of **IAM User Sync**.
- Step 3** After IAM synchronization is complete, create and submit a job on the **Jobs** page.
- End

## 5.30 Error "symbol xxx not defined in file libcrypto.so.1.1" Is Displayed During HA Certificate Generation

### Issue

During HA certificate replacement, when a user runs the `sh ${OMS_RUN_PATH}/workspace/ha/module/hacom/script/gen-cert.sh --root-ca --country=CN --state=state --city=city --company=company --organize=organize --common-name=commonname --email=Cluster user mailbox` command to generate `root-ca.crt` and `root-ca.pem` in the `${OMS_RUN_PATH}/workspace0/ha/local/cert` directory on the active management node, the following error message is displayed:

```
openssl: relocation error: openssl: symbol BIO_new_dgram_sctp version OPENSSL_1_1_0 not defined in file libcrypto.so.1.1 with link time referencecreate server private key failed.
```

### Cause Analysis

- OpenSSL may have been changed and is not the default `/usr/bin/openssl`.
- The dynamic library dependency `libcrypto.so.1.1` cannot be found.
- If the cluster version is 3.2.0 or earlier, the environment variable configuration command (for example, `source bigdata_env`) may have been executed or the environment variable may have been changed before the operation. This issue has been resolved in versions later than 3.2.0.

### Procedure

- Step 1** Log in to the active management node using its IP address as user `omm`.
- Step 2** Run the following command to check whether the command output is `/usr/bin/openssl`. If it is not, change OpenSSL to the default `/usr/bin/openssl`.

```
which openssl
```

- Step 3** Run the following command to check the dynamic library dependency:

```
ldd /usr/bin/openssl
```

The command output is as follows:

```
[omm@xxx ~]$ ldd /usr/bin/openssl
linux-vdso.so.1 (0x0000ffffaf016000)
libssl.so.1.1 => /usr/lib64/libssl.so.1.1 (0x0000ffffae7a000)
libcrypto.so.1.1 => /usr/lib64/libcrypto.so.1.1 (0x0000ffffaebc2000)
libz.so.1 => /usr/lib64/libz.so.1 (0x0000ffffaeb91000)
libdl.so.2 => /usr/lib64/libdl.so.2 (0x0000ffffaeb70000)
libpthread.so.0 => /usr/lib64/libpthread.so.0 (0x0000ffffaeb3b000)
```

```
libc.so.6 => /usr/lib64/libc.so.6 (0x0000ffffae9c5000)
/lib/ld-linux-aarch64.so.1 (0x0000ffffafe8000)
```

Check whether **libcrypto.so.1.1** in the command output points to a value. If **not found** is displayed, run the following command to load **libcrypto.so.1.1**:

**echo \$LD\_LIBRARY\_PATH**

**Step 4** Check whether a non-system OpenSSL library is loaded in the system library environment variables. If yes, change it to the default OpenSSL library of the system.

**Step 5** If the fault persists, contact technical support.

----End

## 5.31 Some Instances Fail to Be Started After Core Nodes Are Added to the MRS Cluster

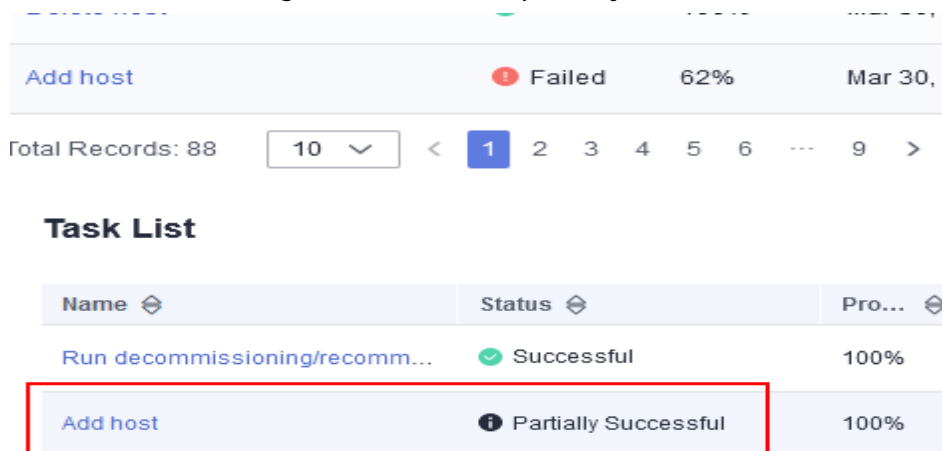
### Symptom

Core nodes are added, but some instances on the nodes may fail to be started. The symptoms are as follows:

1. A core node has been added and is displayed on the **Nodes** page.



2. Some tasks for adding nodes fail or are partially successful.




3. If IAM users have been synchronized, you can view unstarted roles on the **Components** page.

4. If they are not synchronized, you can view unstarted roles on the Manager page of this cluster.

## Procedure

Scenario 1: The task for adding nodes fails before component installation.

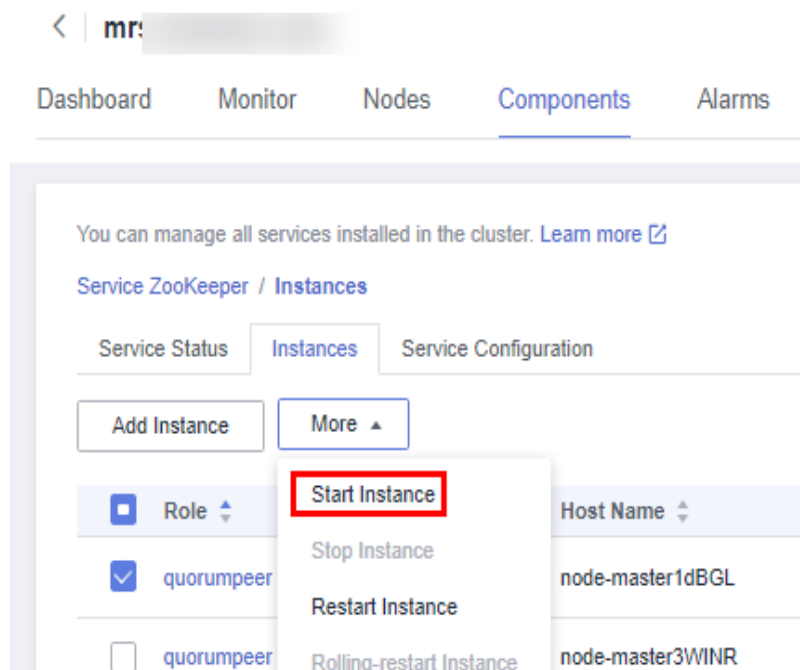
- Step 1** Perform the following steps if the MRS cluster is a pay-per-use cluster:
1. Log in to the MRS console.
  2. Choose **Active Clusters** and click the cluster name to go to the cluster details page.
  3. Click  in the upper part of the page. In the **Task List** column, click the task for adding core nodes.
  4. Records all nodes in the verification request parameter.
  5. Click the **Nodes** tab, select the nodes recorded in [Step 1.4](#), click **Stop** in the upper right corner, and stop the nodes as prompted.
  6. Reduce nodes by referring to [Scaling In a Cluster](#).
- Step 2** If the MRS cluster is billed on a yearly/monthly basis, unsubscribe from the abnormal nodes by referring to [Unsubscribing from a Specified Node in a Yearly/Monthly Cluster](#).

----End

Scenario 2: The task for adding nodes fails after component installation.

- Step 1** Log in to the MRS console.
- Step 2** Choose **Active Clusters** and click the 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.
- Step 4** Click **Components** and check the role status of each service. If a role is not started, select the role, click **More**, and select **Start Instance** to start the instance.





**Step 5** If the startup fails, rectify the fault based on the error information in the task list and try again.

**NOTE**

- If there are many abnormal roles, click **Management Operations** in the upper right corner to start all components.
- For other exceptions that cannot be resolved, contact technical support.
- You can also start the instance on the Manager page of the cluster. For details, see [Overview](#)

----End

# 6 Using Alluixo

## 6.1 Error Message "Does not contain a valid host:port authority" Is Reported When Alluixo Is in HA Mode

### Issue

Error message "Does not contain a valid host:port authority" is reported for Alluixo in HA mode in a security cluster.

### Symptom

Error message "Does not contain a valid host:port authority" is reported for Alluixo in HA mode in a security cluster.

```
java.lang.IllegalArgumentException: Does not contain a valid host:port authority: mode=ana-core@10.5af190a0-ae57-4792-b37c-ef1007558268.com:19998_mode-master2j1ly.5af190a0-ae57-4792-b37c-ef1007558268.com:19998_mode-master2j1ly.5af190a0-ae57-4792-b37c-ef1007558268.com:19998
at org.apache.hadoop.net.NetUtils.createSocketAddr(NetUtils.java:211)
at org.apache.hadoop.net.NetUtils.createSocketAddr(NetUtils.java:164)
at org.apache.hadoop.security.SecurityUtil.buildDTServiceName(SecurityUtil.java:307)
at org.apache.hadoop.fs.FileSystem.getAnonymousServiceName(FileSystem.java:321)
at org.apache.hadoop.fs.FileSystem.getAnonymousServiceName(FileSystem.java:348)
at org.apache.hadoop.fs.FileSystem.addDelegatingTokenIfLegacyToken(FileSystem.java:527)
at org.apache.hadoop.mapreduce.security.TokenCache.obtainTokensForAllModesInternal(TokenCache.java:126)
at org.apache.hadoop.mapreduce.security.TokenCache.obtainTokensForAllModesInternal(TokenCache.java:166)
at org.apache.hadoop.mapreduce.security.TokenCache.obtainTokensForAllModes(TokenCache.java:80)
at org.apache.hadoop.examples.TeraSort.TeraSortOutputFormat.checkOutputTypes(TeraSortOutputFormat.java:98)
at org.apache.hadoop.mapreduce.JobSubmitter.checkOutputs(JobSubmitter.java:268)
at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:141)
at org.apache.hadoop.mapreduce.Job.submitJob(Job.java:1341)
at org.apache.hadoop.mapreduce.Job.submitJob(Job.java:1338)
at java.security.AccessController.doPrivileged(Native Method)
at java.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1848)
at org.apache.hadoop.mapreduce.Job.submitJob(Job.java:1338)
at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1306)
at org.apache.hadoop.examples.TeraSort.TeraSort.run(TeraSort.java:301)
at org.apache.hadoop.examples.TeraSort.TeraSort.main(TeraSort.java:78)
at org.apache.hadoop.examples.TeraSort.TeraSort.main(TeraSort.java:305)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at org.apache.hadoop.util.ProgramDriver$ProgramDescription.invoke(ProgramDriver.java:71)
at org.apache.hadoop.util.ProgramDriver.run(ProgramDriver.java:144)
at org.apache.hadoop.examples.ExampleDriver.main(ExampleDriver.java:74)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at org.apache.hadoop.util.RunJar.main(RunJar.java:131)
```

### Cause Analysis

**org.apache.hadoop.security.SecurityUtil.buildDTServiceName** does not support multiple alluxiomaster addresses in the URI.

### Procedure

Use **alluixo:///** or **alluixo://<IP address or hostname of the active AlluixoMaster>:19998/** for access.

# 7 Using ClickHouse

## 7.1 ClickHouse Fails to Start Due to Incorrect Data in ZooKeeper

### Symptom

An instance node in the ClickHouse cluster fails to start. The startup log of the instance node contains error information similar to the following:

```
2021.03.15 21:01:19.816593 [ 11111 ] {} <Error> Application: DB::Exception:
The local set of parts of table DEFAULT.lineorder doesn't look like the set of
parts in ZooKeeper: 59.99 million rows of 59.99 million total rows in
filesystem are suspicious. There are 30 unexpected parts with 59986052 rows
(14 of them is not just-written with 59986052 rows), 0 missing parts (with 0
blocks):: Cannot attach table `DEFAULT`.`lineorder` from metadata file
...
: while loading database
```

### Cause Analysis

When a ClickHouse instance is abnormal, the ReplicatedMergeTree engine table is repeatedly created in the cluster, and then deleted. The creation and deletion of the ReplicatedMergeTree engine table causes data errors in ZooKeeper, which results in a start failure of ClickHouse.

### Solution

**Step 1** Back up all table data in the database of the faulty node to another directory.

- Back up table data:

MRS 3.0.5 or earlier:

```
cd /srv/BigData/data1/clickhouse/data/Database name
```

```
mkdir -p Backup directory/data1
```

```
mv {Table name} Backup directory/data1/
```

MRS 3.1.0 or later:

```
head -1 /srv/BigData/data1/clickhouse_path/metadata/Database name/Table name.sql | awk -F ' ' '{print $5}' | sed "s/!/g" Target UUID
```

```
cd /srv/BigData/data1/clickhouse/store/{First three characters of the UUID}
mkdir -p Backup directory/data1
mv {UUID} Backup directory/data1/
```

**NOTE**

If there are multiple disks, back up data of **data1** to **dataN**.

- Back up metadata information:

```
cd /srv/BigData/data1/clickhouse_path/metadata/Database name
mv Table name.sql Backup directory
```

For example, to back up table **lineorder** in the **default** database to the **/home/backup** directory, run the following commands:

```
cd /srv/BigData/data1/clickhouse/data/default
mkdir -p /home/backup/data1
mv lineorder /home/backup/data1
cd /srv/BigData/data1/clickhouse_path/metadata
mv lineorder.sql /home/backup
```

**Step 2** Log in to MRS Manager, choose **Cluster > Services > ClickHouse > Instance**, select the target instance node, and click **Start Instance**.

**Step 3** After the instance is started, use the ClickHouse client to log in to the faulty node. There can be security risks if a command contains the authentication password. You are advised to disable the command recording function (history) before running the command.

```
clickhouse client --host ClickHouse instance IP address --user Username --password Password
```

**Step 4** Run the following commands to obtain the ZooKeeper path **zookeeper\_path** of the current table and **replica\_num** of the corresponding node:

```
SELECT zookeeper_path FROM system.replicas WHERE database = 'Database name' AND table = 'Table name';
SELECT replica_num,host_name FROM system.clusters;
```

**Step 5** Run the following command to access the ZooKeeper CLI:

```
zkCli.sh -server IP address of the ZooKeeper node:2181
```

For details about how to obtain the ZooKeeper IP address, see [How Do I Obtain the ZooKeeper Address?](#)

**Step 6** Locate the ZooKeeper path corresponding to the table data of the faulty node.

```
ls zookeeper_path/replicas/replica_num
```

**NOTE**

*zookeeper\_path* indicates the **zookeeper\_path** value obtained in [Step 4](#).

*replica\_num* indicates the **replica\_num** value obtained in [Step 4](#).

**Step 7** Run the following command to delete the replica data from ZooKeeper:

```
deleteall zookeeper_path/replicas/replica_num
```

**Step 8** Use the ClickHouse client to log in to the faulty node and create the ReplicatedMergeTree engine table of the cluster.

```
clickhouse client --host ClickHouse instance IP address --multiline --user  
Username --password Password
```

```
CREATE TABLE Database name.Table name ON CLUSTER Cluster name
```

...

```
ENGINE = ReplicatedMergeTree ...
```

The following error message is displayed on other replica nodes, which is normal and can be ignored:

```
Received exception from server (version 20.8.7):  
Code: 57. DB::Exception: Received from x.x.x.x:9000. DB::Exception:  
There was an error on [x.x.x.x:9000]: Code: 57, e.displayText() =  
DB::Exception: Table DEFAULT.lineorder already exists. (version 20.8.11.17  
(official build)).
```

After the table is successfully created, the table data on the faulty node will be automatically synchronized. The data restoration is complete.

----End

## 7.2 An Exception Occurs When ClickHouse Consumes Kafka Data

### Symptom

A user creates the Kafka engine table **test.user\_log\_kafka** in the ClickHouse cluster to consume Kafka data. However, the Kafka monitoring result shows that messages have been stacked since the early morning and data has not been consumed.

### Cause Analysis

If ClickHouse encounters an exception in consuming data, messages are stacked in Kafka. You need to check the ClickHouse logs to locate the fault.

1. Log in to the MRS cluster, go to the node where the ClickHouse instance is located, and view the **clickhouse-server.log** file in the **/var/log/Bigdata/clickhouse** directory. The following error information is displayed.

```

2021.07.29 11:33:53.739692 [ 15132 ] {} <Error> void DB::StorageKafka::threadFunc(): Code: 27, e.displayText() = DB::Exception: Cannot parse input: expected "" before: ",8,3.000000,2.000000", ip:"", "is_first_time":1, "is_new_user":0, "latest_active_source": "", "project_id":1, "referrer_scene_id": "", "referrer_scene_url": "", "scene_id": (while read the value of key <name_id>: (at row 1)
. Stack trace (when copying this message, always include the lines below):
0. Poco::Exception: Exception(std:: __1::basic_string<char, std:: __1::char_traits<char>, std:: __1::allocator<char> > const&, int) @ 0x1250e59c in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
1. DB::Exception: Exception(std:: __1::basic_string<char, std:: __1::char_traits<char>, std:: __1::allocator<char> > const&, int) @ 0x862d063 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
2. DB::throwAtAssertionFailed(char const*, DB::ReadBuffer&) (.cold) @ 0x862d063 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
3. DB::assertChar(char, DB::ReadBuffer&) @ 0xf214649 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight_ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
4. DB::DataTypeNumberBase<char8_t>::deserializeTextJSON(DB::IColumn&, DB::ReadBuffer&, DB::FormatSettings const&) @ 0xf3527fc in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
5. DB::JSONEachRowRowInputFormat::readField(unsigned long, std:: __1::vector<COW-DB::IColumn>::mutable_ptr<DB::IColumn>, std:: __1::allocator<COW-DB::IColumn>::mutable_ptr<DB::IColumn> > &5) @ 0xf558c1a in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
6. DB::JSONEachRowRowInputFormat::readJSONObject(std:: __1::vector<COW-DB::IColumn>::mutable_ptr<DB::IColumn>, std:: __1::allocator<COW-DB::IColumn>::mutable_ptr<DB::IColumn> > &6) @ 0xf558e29 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
7. DB::JSONEachRowRowInputFormat::readRow(std:: __1::vector<COW-DB::IColumn>::mutable_ptr<DB::IColumn>, std:: __1::allocator<COW-DB::IColumn>::mutable_ptr<DB::IColumn> > &7, DB::RowReadExtension&) @ 0xf559d95 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
8. DB::IRowInputFormat::generate() @ 0x1007b959 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse
9. DB::ISource::work() @ 0xfab037 in /opt/Bigdata/FusionInsight_ClickHouse_8.0.2.1/install/FusionInsight-ClickHouse-v20.8.7.15-lts/clickhouse/bin/clickhouse

```

- The same error log is found on other ClickHouse nodes. Therefore, Kafka messages are stacked because an exception occurs when ClickHouse parses Kafka data.

## Procedure

Run the following command to modify the **kafka\_skip\_broken\_messages** attribute of the table:

```
# ALTER test.user_log MODIFY SETTINGS kafka_skip_broken_messages=10000
```

### NOTE

- Value **10000** in this command can be adjusted based on the proportion of dirty data.
- kafka\_skip\_broken\_messages** indicates the tolerance of the Kafka message parser to schema-incompatible messages per block. The default value is **0**.

For example, if **kafka\_skip\_broken\_messages** is set to *N*, the engine skips *N* Kafka messages that cannot be parsed.

# 8 Using DBService

## 8.1 DBServer Instance Is in Abnormal Status

### Symptom

A DBServer instance is in the **Concerning** state for a long period of time.

**Figure 8-1** DBServer instance status

Role	Host Name	OM IP Address	Business IP Address	Rack	Operating Status	Health Status
<input type="checkbox"/> DBServer(Active)	node-master2iMW	192.168.0.13	192.168.0.13	/default/rack4b34	Started	Good
<input checked="" type="checkbox"/> DBServer(Standby)	node-master1GZBS	192.168.0.53	192.168.0.53	/default/rack4b34	Started	Recovering

### Cause Analysis

The permission for files or directories in the data directory is incorrect. GaussDB requires that the file permission be at least 600 and directory permission be at least 700.

**Figure 8-2** Directory permission list

```
omm@ 192-168-234-176: /srv/BigData/dbdata_service> ll
total 4
drwx----- 19 omm wheel 4096 Dec 14 10:15 data
```

**Figure 8-3** File permission list

```
omm@ 192-168-234-176: /srv/BigData/dbdata_service/data> ll
total 128
drwx----- 6 omm wheel 4096 Dec 9 15:47 base
-rw----- 1 omm wheel 922 Dec 9 15:34 dblink.conf
-rw----- 1 omm wheel 16 Dec 14 10:15 gaussdb.state
drwx----- 2 omm wheel 4096 Dec 14 10:17 global
drwx----- 2 omm wheel 4096 Dec 11 00:00 pg_audit
drwx----- 2 omm wheel 4096 Dec 14 10:15 pg_blackbox
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_clog
drwx----- 2 omm wheel 4096 Dec 14 10:15 pg_confdir_backup
-rw----- 1 omm wheel 1024 Dec 9 15:34 pg_ctl.lock
-rw----- 1 omm wheel 4245 Dec 9 15:47 pg_hba.conf
-rw----- 1 omm wheel 1024 Dec 9 15:47 pg_hba.conf.lock
-rw----- 1 omm wheel 1636 Dec 9 15:34 pg_ident.conf
drwx----- 2 omm wheel 4096 Dec 9 15:38 pg_log
drwx----- 4 omm wheel 4096 Dec 9 15:34 pg_multixact
drwx----- 2 omm wheel 4096 Dec 14 10:15 pg_notify
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_serial
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_snapshots
drwx----- 2 omm wheel 4096 Dec 14 11:56 pg_stat_tmp
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_subtrans
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_tblspc
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_twophase
-rw----- 1 omm wheel 4 Dec 9 15:34 PG_VERSION
drwx----- 2 omm wheel 4096 Dec 9 15:34 pg_wallet
drwx----- 3 omm wheel 4096 Dec 9 15:39 pg_xlog
-rw----- 1 omm wheel 13309 Dec 14 10:15 postgresql.conf
-rw----- 1 omm wheel 1024 Dec 9 15:34 postgresql.conf.lock
-rw----- 1 omm wheel 105 Dec 14 10:15 postmaster.opts
-rw----- 1 omm wheel 96 Dec 14 10:15 postmaster.pid
```

## Solution

- Step 1** Modify the permissions on the files and directories based on the permission list in [Figure 8-2](#) and [Figure 8-3](#).
- Step 2** Restart the DBServer instance.

----End

## 8.2 DBServer Instance Remains in the Restoring State

### Symptom

A DBServer instance remains in the **Restoring** state. The status cannot be recovered even after a restart.

### Cause Analysis

- DBService monitors the ``${BIGDATA_HOME}/MRS_XXX/install/dbservice/ha/module/harm/plugin/script/gSDB/startGS.fail` file. `XXX` indicates the product version.



2. If the value in the file is greater than 3, the startup fails. The NodeAgent keeps trying to restart the instance. In this case, the startup still fails and the value is incremented by 1 each time the startup fails.

## Solution

- Step 1** Log in to MRS Manager.
  - Step 2** Stop the DBServer instance.
  - Step 3** Log in to the node where the DBServer instance is abnormal as user **omm**.
  - Step 4** Change the value of in the `${BIGDATA_HOME}/MRS_XXX/install/dbservice/ha/module/harm/plugin/script/gSDB/.startGS.fail` file to **0**. *XXX* indicates the product version.
  - Step 5** Start the DBServer instance.
- End

## 8.3 Default Port 20050 or 20051 of DBService Is Occupied

### Symptom

DBService restart fails, and information indicating that port 20050 or 20051 is occupied is displayed in the printed fault log.

### Cause Analysis

1. The default port 20050 or 20051 used by DBService is occupied by another process.
2. The DBService process is not stopped, and the port used by DBService is not released.

### Solution

This solution uses port 20051 as an example. The solution to the problem that port 20050 is occupied is similar.

- Step 1** Log in to the node where the error is reported as user **root**, and run the `netstat -nap | grep 20051` command to check the process that occupies port 20051.
- Step 2** Run the `kill` command to forcibly stop the process that uses port 20051.
- Step 3** About 2 minutes later, run the `netstat -nap | grep 20051` command again to check whether any process uses the port.
- Step 4** Check the service to which the process belongs and change the port for the service.
- Step 5** Run the `find . -name "*20051*"` command in the `/tmp` and `/var/run/MRS-DBService/` directories, and delete all files found.

**Step 6** Log in to Manager and restart DBService.

----End

## 8.4 DBServer Instance Is Always in the Restoring State Because the Incorrect /tmp Directory Permission

### Symptom

A DBServer instance remains in the **Restoring** state. The status cannot be recovered even after a restart.

### Cause Analysis

1. Check `/var/log/Bigdata/dbservice/healthCheck/dbservice_processCheck.log`. It is found that GaussDB is abnormal.

**Figure 8-4** GaussDB exception

```
[2019-07-22 10:57:00] ERROR: [:108]: Host 192.168.5.42 gaussdb status is Exception.
[2019-07-22 10:57:00] ERROR: [:154]: Check DBService health failed.
[2019-07-22 10:57:10] INFO: [:84]: check host:192.168.5.42 DBService health.
[2019-07-22 10:57:10] INFO: [:99]: Host 192.168.5.42 floatip status is Normal
Normal.
[2019-07-22 10:57:10] ERROR: [:108]: Host 192.168.5.42 gaussdb status is Exception.
[2019-07-22 10:57:10] ERROR: [:154]: Check DBService health failed.
[2019-07-22 10:57:20] INFO: [:84]: check host:192.168.5.42 DBService health.
[2019-07-22 10:57:20] INFO: [:99]: Host 192.168.5.42 floatip status is Normal
Normal.
[2019-07-22 10:57:20] ERROR: [:108]: Host 192.168.5.42 gaussdb status is Exception.
[2019-07-22 10:57:20] ERROR: [:154]: Check DBService health failed.
[2019-07-22 10:57:30] INFO: [:84]: check host:192.168.5.42 DBService health.
[2019-07-22 10:57:31] INFO: [:99]: Host 192.168.5.42 floatip status is Normal
Normal.
[2019-07-22 10:57:31] ERROR: [:108]: Host 192.168.5.42 gaussdb status is Exception.
[2019-07-22 10:57:31] ERROR: [:154]: Check DBService health failed.
[2019-07-22 10:57:41] INFO: [:84]: check host:192.168.5.42 DBService health.
[2019-07-22 10:57:41] INFO: [:99]: Host 192.168.5.42 floatip status is Normal
```

2. The check result shows that the permission on the `/tmp` directory is incorrect.

**Figure 8-5** /tmp permission

```
[root@node-master1DEdJ DB]# ll / -rlth
total 76K
drwxr-xr-x. 2 root root 4.0K Dec 12 2016 mnt
drwxr-xr-x. 2 root root 4.0K Dec 12 2016 media
drwxr-xr-x. 13 root root 4.0K Jul 15 16:25 usr
-rwxr-xr-x. 1 root root 3.8K Jul 15 16:25 README
-rwxr-xr-x. 1 root root 0 Jul 15 16:25 OTC_EulerOS_2.x86_64-0.9.1-20170904-0513
lrwxrwxrwx. 1 root root 8 Jul 15 16:26 sbin -> usr/sbin
lrwxrwxrwx. 1 root root 9 Jul 15 16:26 lib64 -> usr/lib64
lrwxrwxrwx. 1 root root 7 Jul 15 16:26 lib -> usr/lib
lrwxrwxrwx. 1 root root 7 Jul 15 16:26 bin -> usr/bin
drwxr-xr-x. 3 root root 4.0K Jul 15 16:29 srv
drwxr-xr-x. 7 root root 4.0K Jul 15 16:39 CloudResetPwdUpdateAgent
drwxr-xr-x. 7 root root 4.0K Jul 15 16:39 CloudrResetPwdAgent
drwx----- 2 root root 16K Jul 15 16:46 lost+found
dr-xr-xr-x. 236 root root 0 Jul 19 17:36 proc
dr-xr-xr-x. 4 root root 4.0K Jul 19 17:37 boot
dr-xr-xr-x. 13 root root 0 Jul 19 17:37 sys
drwxr-xr-x. 19 root root 4.0K Jul 19 17:37 var
drwxr-xr-x. 19 root root 3.0K Jul 19 17:37 dev
drwxr-xr-x. 2 root root 4.0K Jul 19 17:38 tmpdir
drwxr-xr-x. 7 root root 4.0K Jul 19 17:38 opt
-rw----- 1 root root 0 Jul 19 17:39 install_os_optimization.log
drwxr-xr-x. 6 root root 4.0K Jul 19 17:54 home
drwxr-xr-x. 86 root root 4.0K Jul 19 17:54 etc
drwxr-xr-x. 30 root root 960 Jul 22 10:49 run
drwx----- 23 root root 4.0K Jul 22 11:42 tmp
drwx----- 5 root root 4.0K Jul 22 11:50 root
```

## Solution

**Step 1** Run the following command to modify the `/tmp` permission:

```
chmod 1777 /tmp
```

**Step 2** Wait until the instance status recovers.

----End

## 8.5 Failed to Execute a DBService Backup Task

### Symptom

No DBService backup file exists in the backup file path.

```
ls /srv/BigData/LocalBackup/default_20190720222358/ -rlth
```

Figure 8-6 Checking the backup file

```
drwx-----. 2 omm wheel 4096 Aug 5 09:00 LdapServer_20190805090027
drwx-----. 2 omm wheel 4096 Aug 5 10:00 LdapServer_20190805100027
drwx-----. 2 omm wheel 4096 Aug 5 09:00 NameNode_20190805090027
drwx-----. 2 omm wheel 4096 Aug 5 10:00 NameNode_20190805100027
drwx-----. 2 omm wheel 4096 Aug 5 09:01 OMS_20190805090027
drwx-----. 2 omm wheel 4096 Aug 5 10:01 OMS_20190805100027
```

### Cause Analysis

- Check the backup log of DBService in `/var/log/Bigdata/dbservice/scriptlog/backup.log`. It is found that the backup is successful but fails to be uploaded to the OMS node.

```
2017-05-18 02:00:54] INFO: [dbservice_backup.sh:528]: Backup file had been saved to V100R002C003PC205_DBSERVICE_20170518020051.tar.gz
2017-05-18 02:00:54] DEBUG: [dbservice_backup.sh:570]: uploadScript:/opt/huawei/Bigdata/dbservice3PC200/sbin/scp_upload.sh, omsFloatIP:192.168.1.2,
srcPath:/opt/huawei/Bigdata/dbservice3PC200/bak.
2017-05-18 02:00:54] INFO: [dbservice_backup.sh:587]: Begin to upload file.
Warning: Permanently added [redacted] (ECDSA) to the list of known hosts.
Authorized users only. All activity may be monitored and reported.
[redacted] connect to host [redacted] port 22: Connection refused.
2017-05-18 02:00:55] ERROR: [dbservice_backup.sh:650]: Upload-file(/opt/huawei/Bigdata/dbservice3PC200/bak) failed.
2017-05-18 02:00:55] ERROR: [dbservice_backup.sh:555]: Scp backupfile to oms error.
2017-05-18 02:00:55] ERROR: [dbservice_backup.sh:928]: main: auto backup failed.
2017-05-18 02:00:55] INFO: [dbservice_backup.sh:929]: main: start create flag file.
2017-05-18 02:00:55] INFO: [dbservice_backup.sh:750]: Send Alarm(AlarmID:[27008] Category:[0] LocationInfo:[DBService:DBServer:hadoopclh2]) successful.
1554.1
```

- The failure is caused by the SSH failure.

```
omm@hadoopclh2:/opt/huawei/Bigdata/dbservice3PC200/sbin> ssh hadoopclh1
Warning: Permanently added 'hadoopclh1,[redacted]' (ECDSA) to the list of known hosts.
Authorized users only. All activity may be monitored and reported.
Last login: Thu May 18 20:18:45 2017 from [redacted]
omm@hadoopclh1:~> ssh [redacted]
Warning: Permanently added '[redacted]' (ECDSA) to the list of known hosts.
Authorized users only. All activity may be monitored and reported.
Last login: Mon Apr 10 10:50:23 2017 from [redacted]
omm@hadoopclh2:~> exit
logout
Connection to [redacted] closed.
omm@hadoopclh1:~> ssh [redacted]
ssh: connect to host [redacted] port 22: Connection refused
```

## Solution

**Step 1** If the network is faulty, contact network engineers.

**Step 2** Perform backup operations again after the network fault is rectified.

----End

## 8.6 Components Failed to Connect to DBService in Normal State

### Symptom

Upper-layer components fail to connect to DBService. The DBService component and two instances are normal.

Figure 8-7 DBService status

Role	Host Name	OM IP	Business IP	Rack	Operating Status	Health Status	Configuration Status
DBServer(Active)	169-120-85-102			jobfaultnac10	Started	Good	Synchronized
DBServer(Standby)	169-120-85-141			jobfaultnac10	Started	Good	Synchronized

### Cause Analysis

1. The upper-layer component is DBService connected through **dbservice.floatip**.
2. Run the **netstat -anp | grep 20051** command on the node where DBServer resides. It is found that the Gauss process of DBService is not bound to the floating IP address during startup, and only local IP address 127.0.0.1 is monitored.

### Solution

**Step 1** Restart the DBService service.

**Step 2** Run the **netstat -anp | grep 20051** command on the active DBServer node to check whether **dbservice.floatip** is bound.

----End

## 8.7 DBServer Failed to Start

### Symptom

DBService fails to be started and restarts also fail. The instance keeps in the **Recovering** state.

Figure 8-8 DBService status

Role	Host Name	OM IP Address	Business IP Address	Rack	Operating Status	Health Status
DBServer(Active)	node-master2iM1W	192.168.0.13	192.168.0.13	/default/rack4b34	Started	Good
DBServer(Standby)	node-master1GZ8S	192.168.0.53	192.168.0.53	/default/rack4b34	Started	Recovering

### Cause Analysis

1. Check the DBService logs in **/var/log/Bigdata/dbservice/DB/gs\_ctl-current.log**. The following error message is displayed:

```

OCCATION: PostmasterMain, postmaster.c:798
OO: Starting SelectConfigFiles (postmaster.c:1049)
2017-09-23 15:19:03.591 CST] gaussmaster 922216 LOG: Starting checkDataDir (postmaster.c:1068)
2017-09-23 15:19:03.591 CST] gaussmaster 922216 LOG: Starting ChangeToDataDir (postmaster.c:1074)
2017-09-23 15:19:03.591 CST] gaussmaster 922216 LOG: Starting checkBaseTables (postmaster.c:1120)
2017-09-23 15:19:03.591 CST] gaussmaster 922216 LOG: Starting CreateBarabirLockFile (postmaster.c:1151)
2017-09-23 15:19:03.596 CST] gaussmaster 922216 LOG: Starting pgaudit agent_init (postmaster.c:1169)
2017-09-23 15:19:03.596 CST] gaussmaster 922216 LOG: Starting process_shared_preload_libraries (postmaster.c:1178)
2017-09-23 15:19:03.597 CST] gaussmaster 922216 LOG: could not bind IPv4 socket at the 0 time: ?????????? (pgcomm.c:562)
2017-09-23 15:19:03.597 CST] gaussmaster 922216 HINT: Is another postmaster already running on port 20051? If not, wait a few seconds and retry.
2017-09-23 15:19:03.698 CST] gaussmaster 922216 LOG: could not bind IPv4 socket at the 1 time: ?????????? (pgcomm.c:562)
2017-09-23 15:19:03.698 CST] gaussmaster 922216 HINT: Is another postmaster already running on port 20051? If not, wait a few seconds and retry.
2017-09-23 15:19:03.798 CST] gaussmaster 922216 LOG: could not bind IPv4 socket at the 2 time: ?????????? (pgcomm.c:562)
2017-09-23 15:19:03.798 CST] gaussmaster 922216 HINT: Is another postmaster already running on port 20051? If not, wait a few seconds and retry.
2017-09-23 15:19:03.898 CST] gaussmaster 922216 WARNING: could not create listen socket for "192.168.5.162" (postmaster.c:1235)
2017-09-23 15:19:03.898 CST] gaussmaster 922216 LOG: discard audit data: could not create lock file "/tmp/.s.PGSQL.20051.lock": ??? (pgaudit.c:1961)
2017-09-23 15:19:03.898 CST] gaussmaster 922216 FATAL: could not create lock file "/tmp/.s.PGSQL.20051.lock": ??? (miscinit.c:854)
    
```

2. It is found that the **/tmp** permission is incorrect. The correct value should be **777**.

```

vmr@hadoopc1h2:/var/log/Bigdata/dbservice/DB> ll /
total 100
-rwxr-xr-x  2 root  root    4096 Aug  6  2016 bin
-rwxr-xr-x  3 root  root    4096 Aug  6  2016 boot
-rwxr-xr-x 17 root  root    5080 Sep 20 11:30 dev
-rwxr-xr-x  3 httpd common  0 Sep 20 11:20 ecramafs
-rwxr-xr-x 71 root  root    4096 Sep 22 02:40 etc
-rw-r----- 1 root  root      0 Sep 11 08:25 fsck_corrected_
-rwxr-xr-x  9 root  root    4096 Sep 18 14:39 home
-rwxr-xr-x 12 root  root    4096 Sep 14  2016 lib
-rwxr-xr-x  8 root  root   12288 Sep 14  2016 lib64
-rwx-----  2 root  root   16384 Aug  7  2016 lost+found
-rwxr-xr-x  2 root  root    4096 May  5  2010 media
-rwxr-xr-x  2 root  root    4096 May  5  2010 mnt
-rwxr-xr-x 19 root  root    4096 Jun 30 10:04 opt
-r-xr-xr-x 424 root  root      0 Sep 20 19:18 proc
-rwx-----  5 root  root    4096 Sep 23 10:21 root
-rwxrwxr-x  4 root  root    4096 Aug  7  2016 rrdtool
-rwxr-xr-x  3 root  root   12288 Sep 14  2016 sbin
-rwxr-xr-x  2 root  root    4096 May  5  2010 selinux
-rwxrwxrwx 10 root  root    4096 Nov 15  2016 srv
-rwxr-xr-x 12 root  root      0 Sep 20 11:19 sys
-rwxrwxrwx  1 root  root      1 Aug  7  2016 target -> /
-rwxr-xr-x  6 root  root    4096 Sep 23 15:19 tmp
-rwxr-xr-x 13 root  root    4096 Apr 22  2014 usr
    
```

## Solution

- Step 1 Modify the **/tmp** permission by changing the value to **777**.
  - Step 2 Restart DBService.
- End

## 8.8 DBService Backup Failed Because the Floating IP Address Is Unreachable

### Symptom

The default DBService backup fails, but backups of NameNode, LdapServer, and OMS are successful.

### Cause Analysis

1. Check the error information on the DBService backup page:  
Clear temporary files at backup checkpoint DBService\_test\_DBService\_DBService\_20180326155921 that failed last time.

Temporary files at backup checkpoint DBService\_test\_DBService\_DBService20180326155921 that failed last time are cleared successfully.

```
Start executing the backup task.
The backup of configuration DBService is started.
Check the backup available disk space.
Backup initialization succeeded for configuration DBService.
Clear temporary files at backup checkpoint DBService_test_DBService_DBService_20180326155921 that failed last time.
Temporary files at backup checkpoint DBService_test_DBService_DBService_20180326155921 that failed last time are cleared successfully.
Checkpoint DBService_test_DBService_DBService_20180326162235 is verified successfully before backup.
Temporary files are cleared successfully before backup checkpoint DBService_test_DBService_DBService_20180326162235.
Prestart backup succeeded for checkpoint DBService_test_DBService_DBService_20180326162235.
The snapshot is created successfully for checkpoint DBService_test_DBService_DBService_20180326162235 before backup.
Backup is being performed for checkpoint DBService_test_DBService_DBService_20180326162235.
Backup execution failed. Task ID: 2
Detail: DBService backup task failed, please view details in logs.
Temporary files are cleared successfully after backup checkpoint DBService_test_DBService_DBService_20180326162235.
checkpoint DBService_test_DBService_DBService_20180326162235 is deleted successfully after backup failure.
Failed to backup configuration DBService.
```

2. Check the `/var/log/Bigdata/dbservice/scriptlog/backup.log` file. It is found that the log printing stops and no related backup information is found.
3. Check the `/var/log/Bigdata/controller/backupplugin.log` file on the active OMS node. The following error information is found:  
result error is ssh:connect to host 172.16.4.200 port 22: Connection refused (172.16.4.200 is the floating IP address of DBService)  
DBService backup failed.

```
2018-03-27 07:00:35,758 INFO [pool-1-thread-5] Create adapter from com.huawei.bigdata.cm.backup.MetadataPluginAdapter success.
com.huawei.bigdata.cm.backup.plugin.AbstractBackupRecoveryPlugin.initializePluginAdapter(AbstractBackupRecoveryPlugin.java:92)
2018-03-27 07:00:35,759 INFO [pool-1-thread-5] floatIp is 172.16.4.200. com.huawei.bigdata.cm.db.service.backup.BackupRecoveryPlugin.getFloatIp(BackupRecoveryPlugin.java:233)
2018-03-27 07:00:35,759 INFO [pool-1-thread-5] cmd is ssh 172.16.4.200 /opt/huawei/Bigdata/FusionInsight_V100R002C60020/dbservice/sbin/dbservice_backup.sh -b -d
/srv/BigData/LocalBackup/default_20180326213206/DBService_20180327070010. com.huawei.bigdata.cm.db.service.backup.BackupRecoveryPlugin.startBackup(BackupRecoveryPlugin.java:166)
2018-03-27 07:00:35,759 INFO [pool-1-thread-5] create task taskId is 6. com.huawei.bigdata.cm.db.service.backup.BackupRecoveryPlugin.startBackup(BackupRecoveryPlugin.java:169)
2018-03-27 07:00:35,760 INFO [pool-1-thread-5] startBackup result OperateResult{errorCode:RUNNING, result:6, detailInfo:, packageName:null}.
com.huawei.bigdata.cm.backup.BackupPluginContainerHandler.startBackup(BackupPluginContainerHandler.java:246)
2018-03-27 07:00:35,760 INFO [Thread-132] Executing the command with arguments and env, timeout: 900000
com.huawei.bigdata.cm.controller.api.extern.monitor.script.LinuxScriptExecutionHandler.logMessage(LinuxScriptExecutionHandler.java:64)
2018-03-27 07:00:35,863 INFO [Thread-132] Execute command : /opt/huawei/Bigdata/cm-0.0.1/sbin/scriptlauncher.sh ssh 172.16.4.200
/opt/huawei/Bigdata/FusionInsight_V100R002C60020/dbservice/sbin/dbservice_backup.sh -b -d /srv/BigData/LocalBackup/default_20180326213206/DBService_20180327070010.
com.huawei.bigdata.cm.db.service.backup.BackupTask.run(BackupTask.java:48)
2018-03-27 07:00:35,863 INFO [Thread-132] result status is 255. com.huawei.bigdata.cm.db.service.backup.BackupTask.run(BackupTask.java:49)
2018-03-27 07:00:35,863 INFO [Thread-132] result output is . com.huawei.bigdata.cm.db.service.backup.BackupTask.run(BackupTask.java:50)
2018-03-27 07:00:35,863 INFO [Thread-132] result error is ssh: connect to host 172.16.4.200 port 22: Connection refused
. com.huawei.bigdata.cm.db.service.backup.BackupTask.run(BackupTask.java:51)
2018-03-27 07:00:35,863 ERROR [Thread-132] DBService backup failed. com.huawei.bigdata.cm.db.service.backup.BackupTask.run(BackupTask.java:64)
2018-03-27 07:00:40,868 INFO [pool-1-thread-5] query backup taskId is 6. com.huawei.bigdata.cm.db.service.backup.BackupRecoveryPlugin.getBackupProgress(BackupRecoveryPlugin.java:247)
```

## Solution

- Step 1 Log in to the active DBService node (the Master node bound with the DBService floating IP address).

```
[root@node-master1cuEb ~]# ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.2.223 netmask 255.255.255.0 broadcast 192.168.2.255
    ether fa:16:3e:eb:7e:74 txqueuelen 1000 (Ethernet)
    RX packets 125672126 bytes 35833339919 (33.3 GiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 111023825 bytes 33326544401 (31.0 GiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

eth0:DBS: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.2.206 netmask 255.255.255.0 broadcast 192.168.2.255
    ether fa:16:3e:eb:7e:74 txqueuelen 1000 (Ethernet)

eth0:FI_HUE: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.2.197 netmask 255.255.255.0 broadcast 192.168.2.255
    ether fa:16:3e:eb:7e:74 txqueuelen 1000 (Ethernet)
```

- Step 2 Add the DBService floating IP address to `ListenAddress` or comment out `ListenAddress` in the `/etc/ssh/sshd_config` file.

**Step 3** Run the following command to restart the SSHD service:

```
service sshd restart
```

**Step 4** Check whether the next DBService backup is successful.

----End

## 8.9 DBService Failed to Start Due to the Loss of the DBService Configuration File

### Symptom

The nodes are powered off unexpectedly, and the standby DBService node fails to be restarted.

### Cause Analysis

1. The `/var/log/Bigdata/dbservice/DB/gaussdb.log` file is viewed, which contains no information.
2. The `/var/log/Bigdata/dbservice/scriptlog/preStartDBService.log` file is viewed. This file contains the following information, indicating that the configuration information is lost:  
The program "gaussdb" was found by "`/opt/Bigdata/MRS_XXX/install/dbservice/gaussdb/bin/g_s_guc`"  
But not was not the same version as `g_s_guc`.  
Check your installation.

```
CST 2018-05-07 15:02:09 [ha config]: config runlogpath as /var/log/Bigdata/dbservice already.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:729]: config HA core log: /opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice/ha/module/hacon/script/config_ha.sh -o "/var
CST 2018-05-07 15:02:09 [ha config]: config corepath as /var/log/Bigdata/dbservice/core already.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:729]: config HA script log: /opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice/ha/module/hacon/script/config_ha.sh -k "/v
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:729]: HA log config success.
CST 2018-05-07 15:02:09 [ha config]: config scriptlogpath as /var/log/Bigdata/dbservice already.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:576]: HA config success.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:367]: Finish to config ha server.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:329]: Start to register DBService plugins to HA.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:340]: Finished to register DBService plugins to HA.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:259]: Start modify floatip.xml,g_usFloatIPNetmask:255.255.0.0,g_usGateway:g_usFloatIP:192.168.200.201
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:268]: Finish modify floatip.xml.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:276]: Start modify dbservice_sync.xml,g_dbInstallPath:/opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:276]: Finish modify dbservice_sync.xml.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:813]: Start to copy GaussDBs confs.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:824]: copy GaussDBs confs successfully.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:587]: prestart-dbservice.sh:587:(configGauss)
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:588]: start to config gauss...
[2018-05-07 15:02:09] WARN: [prestart-dbservice.sh:299]: db is not running now. [gs_ctl: no server running].
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:688]: GAUSSDB is not running,return value is 1.
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:618]: start to config gauss end...Execute: /opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice/gaussdb/bin/g_s_guc -D /srv/
ocalhost-192.168.200.197 localport=24056 remotehost=192.168.200.196 remotepart=24056*)
[2018-05-07 15:02:09] INFO: [prestart-dbservice.sh:616]: GAUSSHOME:/opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice/gaussdb;PATH:/opt/haawei/Bigdata/FusionInsight_U100R02
/opt/haawei/Bigdata/jdk-8.0.112:/opt/haawei/Bigdata/jdk1-8.0.112:/opt/haawei/Bigdata/jdk1-8.0.112/bin:/opt/haawei/Bigdata/jdk1-8.0.112/bin:/opt/haawei/Bigdata/jdk1-8.0.112/z
/usr/local/bin:/usr/bin:/usr/sbin:/usr/lib64/bin:/usr/games:--opt/haawei/Bigdata/000-U100R02C6020-x86_64/toolz:/home/omni/keberos/bin;LD_LIBRARY_PATH:/opt/haawei/Bigdata/FusionInsigh
data/000-U100R02C6020-x86_64/lib:/opt/haawei/Bigdata/000-U100R02C6020-x86_64/lib:/opt/haawei/Bigdata/nodeagent/lib;GAUSSD01A:/srv/Bigdata/dbdata_service/data.
The program "gaussdb" was found by "/opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice/gaussdb/bin/g_s_guc"
but was not the same version as g_s_guc.
Check your installation.
[2018-05-07 15:02:09] ERROR: [prestart-dbservice.sh:621]: Gauss config failure,Execute: /opt/haawei/Bigdata/FusionInsight_U100R02C6020/dbservice/gaussdb/bin/g_s_guc -D /srv/Bigdat
s-192.168.200.197 localport=24056 remotehost=192.168.200.196 remotepart=24056*);return 1].
[2018-05-07 15:02:09] ERROR: [prestart-dbservice.sh:916]: Failed to config gauss database.
```

3. The configuration file in the `/srv/BigData/dbdata_service/data` directory on the active DBServer node is compared with the configuration file in the `/srv/BigData/dbdata_service/data` directory on the standby DBServer node, which shows major difference.

```
onn@hadoopc1h3:/srv/BigData/dbdata_service/data> ll
total 128
-rw----- 1 onn wheel    4 May  8 09:54 PG_VERSION
drwx----- 2 onn wheel 4096 May  8 09:54 bak
drwx----- 7 onn wheel 4096 May  8 09:54 base
-rw----- 1 onn wheel  922 May  8 09:54 dblink.conf
-rw----- 1 onn wheel   16 May  8 09:59 gaussdb.state
drwx----- 2 onn wheel 4096 May  8 09:58 global
drwx----- 2 onn wheel 4096 May  8 09:54 pg_audit
drwx----- 2 onn wheel 4096 May  8 09:58 pg_blackbox
drwx----- 2 onn wheel 4096 May  8 09:54 pg_clog
drwx----- 2 onn wheel 4096 May  8 09:58 pg_confdir_backup
-rw----- 1 onn wheel    0 May  8 09:54 pg_ctl.lock
-rw----- 1 onn wheel 4287 May 18 2017 pg_hba.conf
-rw----- 1 onn wheel 1024 May  8 09:54 pg_hba.conf.lock
-rw----- 1 onn wheel 1636 May  8 09:54 pg_ident.conf
drwx----- 2 onn wheel 4096 May  8 09:54 pg_log
drwx----- 4 onn wheel 4096 May  8 09:54 pg_multixact
drwx----- 2 onn wheel 4096 May  8 09:58 pg_notify
drwx----- 2 onn wheel 4096 May  8 09:54 pg_serial
drwx----- 2 onn wheel 4096 May  8 09:54 pg_snapshots
drwx----- 2 onn wheel 4096 May  8 09:58 pg_stat_tmp
drwx----- 2 onn wheel 4096 May  8 09:54 pg_subtrans
drwx----- 2 onn wheel 4096 May  8 09:54 pg_tblspc
drwx----- 2 onn wheel 4096 May  8 09:54 pg_twophase
drwx----- 2 onn wheel 4096 May  8 09:54 pg_wallet
drwx----- 3 onn wheel 4096 May  8 09:54 pg_xlog
-rw----- 1 onn wheel 15277 May  8 09:59 postgresql.conf
-rw----- 1 onn wheel 1024 May  8 09:54 postgresql.conf.lock
-rw----- 1 onn wheel  134 May  8 09:59 postmaster.opts
-rw----- 1 onn wheel  127 May  8 09:58 postmaster.pid
```

```
onn@hadoopc1h3:/srv/BigData/dbdata_service> cd data_bak/
onn@hadoopc1h3:/srv/BigData/dbdata_service/data_bak> ll
total 64
-rw----- 1 onn wheel  202 Feb 11 10:43 backup_label
-rw----- 1 onn wheel    0 Feb 11 10:42 build_completed.start
-rw----- 1 onn wheel   16 Apr 28 17:32 gaussdb.state
-rw----- 1 onn wheel    7 Apr 28 17:32 gs_build.pid
-rwx----- 2 onn wheel 4096 Feb 11 10:44 pg_audit
-rwx----- 2 onn wheel 4096 Feb 11 10:41 pg_blackbox
-rwx----- 2 onn wheel 4096 Feb 11 10:09 pg_confdir_backup
-rw----- 1 onn wheel    0 Apr 28 17:32 pg_ctl.lock
-rw----- 1 onn wheel 4287 May 18 2017 pg_hba.conf
-rwx----- 2 onn wheel 4096 Feb 11 10:43 pg_notify
-rwx----- 2 onn wheel 4096 Feb 11 10:43 pg_xlog
-rw----- 1 onn wheel 15155 May  7 15:33 postgresql.conf
-rw----- 1 onn wheel 1024 May  7 15:33 postgresql.conf.lock
-rw----- 1 onn wheel  134 Feb 11 10:42 postmaster.opts
```

## Solution

- Step 1** Copy the content in the `/srv/BigData/dbdata_service/data` directory on the active node to the standby node and ensure that the file permission and owner group are the same as those on the active node.
- Step 2** Modify configuration in `postgresql.conf`. Set `localhost` to the IP of the local node and `remotehost` to the IP of the peer node.

```
#-----
# CUSTOMIZED OPTIONS
#-----

# Add settings for extensions here
max_files_per_process = 300
unix_socket_directory = '/var/run/FusionInsight-DBService'
replconninfo1 = 'localhost-192.168.200.197 localport=20050 remotehost-192.168.200.196 remoteport=20050'
"postgresql.conf" 382L, 15277C
```



- Step 3** Log in to Manager and restart the standby DBServer node.  
----End

# 9 Using Flink

## 9.1 Error Message "Error While Parsing YAML Configuration File: Security.kerberos.login.keytab" Is Displayed When a Command Is Executed on the Flink Client

### Symptom

The client is successfully installed. When a user runs a client command, for example, `yarn-session.sh`, the following error message is displayed:

```
[root@host01 bin]# yarn-session.sh
2018-10-25 01:22:06,454 | ERROR | [main] | Error while trying to split key and value in configuration
file /opt/flinkclient/Flink/flink/conf/flink-conf.yaml:80: "security.kerberos.login.keytab: " |
org.apache.flink.configuration.GlobalConfiguration (GlobalConfiguration.java:160)
Exception in thread "main" org.apache.flink.configuration.IllegalConfigurationException: Error while
parsing YAML configuration file :80: "security.kerberos.login.keytab: "
```

### Cause Analysis

In a secure cluster environment, Flink requires security authentication. The security authentication is not configured on the current client.

1. The following two authentication modes are available for Flink.
  - Kerberos authentication: Flink Yarn client, Yarn ResourceManager, JobManager, HDFS, TaskManager, Kafka, and ZooKeeper
  - Internal authentication mechanism of Yarn: The internal authentication used between YarnResource Manager and Application Master (AM).
2. If a security cluster is required, the Kerberos authentication and security cookie authentication are mandatory. As shown in the logs, it is found that the **security.kerberos.login.keytab** setting in the configuration file is incorrect and the security configuration is not performed.

## Solution

**Step 1** Download the user **keytab** file from MRS and save it to a directory on the node where the Flink client is deployed.

**Step 2** Configure the **flink-conf.yaml** file.

1. Keytab path

```
security.kerberos.login.keytab: /home/flinkuser/keytab/abc222.keytab
```

### NOTE

- **/home/flinkuser/keytab/abc222.keytab** indicates the user directory, which is the directory saves the keytab file in **Step 1**.
  - Ensure that the client user has the permission on the corresponding directory.
2. Principal name
- ```
security.kerberos.login.principal: abc222
```
3. In HA mode, if Zookeeper is configured, the ZooKeeper Kerberos authentication configuration items must be configured as follows:
- ```
zookeeper.sasl.disable: false  
security.kerberos.login.contexts: Client
```
4. If Kerberos authentication is required between the Kafka client and Kafka broker, configure it as follows:
- ```
security.kerberos.login.contexts: Client,KafkaClient
```

----End

## 9.2 Error Message "Error while parsing YAML configuration file : security.kerberos.login.principal:pippo" Is Displayed When a Command Is Executed on the Flink Client

### Symptom

The client is successfully installed. When a user runs a client command, for example, **yarn-session.sh**, the following error message is displayed:

```
[root@host01 bin]# yarn-session.sh  
2018-10-25 19:27:01,397 | ERROR | [main] | Error while trying to split key and value in configuration  
file /opt/flinkclient/Flink/flink/conf/flink-conf.yaml:81: "security.kerberos.login.principal:pippo " |  
org.apache.flink.configuration.GlobalConfiguration (GlobalConfiguration.java:160)  
Exception in thread "main" org.apache.flink.configuration.IllegalConfigurationException: Error while  
parsing YAML configuration file :81: "security.kerberos.login.principal:pippo "  
    at org.apache.flink.configuration.GlobalConfiguration.loadYAMLResource(GlobalConfiguration.java:161)  
    at org.apache.flink.configuration.GlobalConfiguration.loadConfiguration(GlobalConfiguration.java:112)  
    at org.apache.flink.configuration.GlobalConfiguration.loadConfiguration(GlobalConfiguration.java:79)  
    at org.apache.flink.yarn.cli.FlinkYarnSessionCli.main(FlinkYarnSessionCli.java:482)
```

### Cause Analysis

The format of configuration items in the **flink-conf.yaml** file is incorrect.

```
security.kerberos.login.principal:pippo
```

## Solution

Modify the configurations in the **flink-conf.yaml** file and ensure that there is a space between a configuration item and its value.

```
security.kerberos.login.principal: pippo
```

## 9.3 Error Message "Could Not Connect to the Leading JobManager" Is Displayed When a Command Is Executed on the Flink Client

### Symptom

During the creation of the Flink cluster, the following error message is displayed after the **yarn-session.sh** command execution is suspended for a while:

```
2018-09-20 22:51:16,842 | WARN | [main] | Unable to get ClusterClient status from Application Client |
org.apache.flink.yarn.YarnClusterClient (YarnClusterClient.java:253)
org.apache.flink.util.FlinkException: Could not connect to the leading JobManager. Please check that the
JobManager is running.
  at org.apache.flink.client.program.ClusterClient.getJobManagerGateway(ClusterClient.java:861)
  at org.apache.flink.yarn.YarnClusterClient.getClusterStatus(YarnClusterClient.java:248)
  at org.apache.flink.yarn.YarnClusterClient.waitForClusterToBeReady(YarnClusterClient.java:516)
  at org.apache.flink.yarn.cli.FlinkYarnSessionCli.run(FlinkYarnSessionCli.java:717)
  at org.apache.flink.yarn.cli.FlinkYarnSessionCli$1.call(FlinkYarnSessionCli.java:514)
  at org.apache.flink.yarn.cli.FlinkYarnSessionCli$1.call(FlinkYarnSessionCli.java:511)
  at java.security.AccessController.doPrivileged(Native Method)
  at javax.security.auth.Subject.doAs(Subject.java:422)
  at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1729)
  at org.apache.flink.runtime.security.HadoopSecurityContext.runSecured(HadoopSecurityContext.java:41)
  at org.apache.flink.yarn.cli.FlinkYarnSessionCli.main(FlinkYarnSessionCli.java:511)
Caused by: org.apache.flink.runtime.leaderretrieval.LeaderRetrievalException: Could not retrieve the leader
gateway.
  at org.apache.flink.runtime.util.LeaderRetrievalUtils.retrieveLeaderGateway(LeaderRetrievalUtils.java:79)
  at org.apache.flink.client.program.ClusterClient.getJobManagerGateway(ClusterClient.java:856)
  ... 10 common frames omitted
Caused by: java.util.concurrent.TimeoutException: Futures timed out after [10000 milliseconds]
```

### Possible Causes

The SSL communication encryption is enabled for Flink, but no correct SSL certificate is configured.

### Solution

For MRS 2.x or earlier, perform the following operations:

Method 1:

Run the following command to disable the Flink SSL communication encryption, and modify the client configuration file **conf/flink-conf.yaml**.

```
security.ssl.internal.enabled: false
```

Method 2:

Enable the Flink SSL communication encryption and retain the default value of **security.ssl.internal.enabled**.

Configure the SSL as follows:

- If the **keystore** or **truststore** file path is a relative path, allow the Flink client directory where the command is executed to access this relative path directly.

```
security.ssl.internal.keystore: ssl/flink.keystore  
security.ssl.internal.truststore: ssl/flink.truststore
```

Add **-t** option to the CLI **yarn-session.sh** command of Flink to transmit the KeyStore and TrustStore files to each execution node.

```
yarn-session.sh -t ssl/ 2
```

- If the keystore or truststore file path is an absolute path, the keystore or truststore files must exist in the absolute path on Flink Client and all nodes.

```
security.ssl.internal.keystore: /opt/client/Flink/flink/conf/flink.keystore  
security.ssl.internal.truststore: /opt/client/Flink/flink/conf/flink.truststore
```

For MRS 3.x or later, perform the following operations:

Method 1:

Run the following command to disable the Flink SSL communication encryption, and modify the client configuration file **conf/flink-conf.yaml**.

```
security.ssl.enabled: false
```

Method 2:

Enable the Flink SSL communication encryption and retain the default value of **security.ssl.enabled**.

Configure the SSL as follows:

- If the **keystore** or **truststore** file path is a relative path, allow the Flink client directory where the command is executed to access this relative path directly.

```
security.ssl.keystore: ssl/flink.keystore  
security.ssl.truststore: ssl/flink.truststore
```

Add **-t** option to the CLI **yarn-session.sh** command of Flink to transmit the KeyStore and TrustStore files to each execution node.

```
yarn-session.sh -t ssl/ 2
```

- If the keystore or truststore file path is an absolute path, the keystore or truststore files must exist in the absolute path on Flink Client and all nodes.

```
security.ssl.keystore: /opt/client/Flink/flink/conf/flink.keystore  
security.ssl.truststore: /opt/client/Flink/flink/conf/flink.truststore
```

## 9.4 Failed to Create a Flink Cluster by Running yarn-session As Different Users

### Symptom

Two users **testuser** and **bdpuser** with the same rights are used to create the Flink cluster.

When user **testuser** is used to create a Flink cluster, no error message is displayed. While user **bdpuser** is used to create a Flink cluster, an error message is displayed during the **yarn-session.sh** command execution:

```
2019-01-02 14:28:09,098 | ERROR | [main] | Ensure path threw exception |  
org.apache.flink.shaded.curator.org.apache.curator.framework.impls.CuratorFrameworkImpl  
(CuratorFrameworkImpl.java:566)  
org.apache.flink.shaded.zookeeper.org.apache.zookeeper KeeperException$NoAuthException:  
KeeperErrorCode = NoAuth for /flink/application_1545397824912_0022
```

## Possible Causes

The HA configuration item is not modified.

In the Flink configuration file, the default value of **high-availability.zookeeper.client.acl** is **creator**, indicating that only the creator has the access permission. A new user cannot access the directory on ZooKeeper. As a result, the **yarn-session.sh** command execution fails.

## Solution

**Step 1** Change the value of **high-availability.zookeeper.path.root** in the **conf/flink-conf.yaml** file, for example:

```
high-availability.zookeeper.path.root: flink2
```

**Step 2** Submit the Flink task again.

----End

## 9.5 Flink Service Program Fails to Read Files on the NFS Disk

### Issue

The Flink service program cannot read files on the NFS disk mounted to the cluster node.

### Symptom

The Flink service program developed by a user needs to read the user-defined configuration file. The configuration file is stored on the NFS disk. The NFS disk is mounted to the cluster node and can be accessed by all nodes in the cluster. After the user submits the Flink program, the service code cannot access the user-defined configuration file. As a result, the service program fails to be started.

### Cause Analysis

The root cause is that the permission on the root directory of the NFS disk is insufficient. As a result, the Flink program cannot access the directory after being started.

Flink tasks of MRS are running on Yarn. If Kerberos authentication is not enabled for the cluster, the user who runs the tasks on Yarn is **yarn\_user**. If the user-defined configuration file is used after the tasks are started, **yarn\_user** must be allowed to access the file and the parent directory of the file (parent directory of the file on the NFS, not the soft link on the cluster node). Otherwise, the program cannot obtain the file content. If Kerberos authentication is enabled for the cluster, the file permission must allow the user who submits the program to access the file.

## Procedure

**Step 1** Log in to the master node in the cluster as user **root**.

**Step 2** Run the following command to check the permission on the parent directory of the user-defined configuration file:

```
ll <Parent directory of the file path>
```

**Step 3** Go to the directory of the file to be accessed on the NFS disk and change the permission of the parent directory of the user-defined configuration file to 755.

```
chmod 755 -R /<Parent directory of the file path>
```

**Step 4** Check whether the core or task node can access the configuration file.

1. Log in to the core or task node as user **root**.

If Kerberos authentication is enabled for the current cluster, log in to the core node as user **root**.

2. Run **su - yarn\_user** to switch to user **yarn\_user**.

If Kerberos authentication is enabled for the cluster, run the **su - User who submits the job** command to switch the user.

3. Run the following command to check the user permission. The file path must be the absolute path of the file.

```
ll <File path>
```

----End

## Summary and Suggestions

When a user-defined configuration file needs to be accessed in the submitted task, especially when the NFS disk is mounted, you need to check whether the permission of the parent directory of the file is correct in addition to the file permission. When an NFS disk is mounted to an MRS cluster node, a soft link is created to the NFS directory. In this case, you need to check whether the directory permission on the NFS is correct.

## 9.6 Failed to Customize the Flink Log4j Log Level

### Issue

The customized level for Flink Log4j logs of an MRS 3.1.0 cluster does not take effect.

### Symptom

1. When analyzing data using Flink of an MRS 3.1.0 cluster, a user changes the log level in the **log4j.properties** file in the **\$Flink\_HOME/conf** directory to **INFO**.
2. However, after the task is submitted successfully, the log level displayed on the console is still **ERROR**, rather than **INFO**.

## Cause Analysis

The **log4j.properties** file in the **\$Flink\_HOME/conf** directory controls the log output of in JobManager and TaskManager operators, and the logs are printed to the corresponding Yarn containers. You can view the logs on the Yarn web UI.

In MRS 3.1.0 and later versions, the default log framework of Flink 1.12.0 is Log4j2. The configuration method is different from that of Log4j. For example, Log4j log rules do not take effect.

## Procedure

For details about configuring Log4j2 log specifications, see the official open-source document at <http://logging.apache.org/log4j/2.x/manual/configuration.html#Properties>.



# 10 Using Flume

---

## 10.1 Class Cannot Be Found After Flume Submits Jobs to Spark Streaming

### Issue

After Flume submits jobs to Spark Streaming, the class cannot be found.

### Symptom

After the Spark Streaming code is packed into a JAR file and submitted to the cluster, an error message is displayed indicating that the class cannot be found. The following two methods are not useful:

1. When submitting a Spark job, run the `--jars` command to reference the JAR file of the class.
2. Import the JAR file where the class resides to the JAR file of Spark Streaming.

### Cause Analysis

Some JAR files cannot be loaded during Spark job execution, resulting that the class cannot be found.

### Procedure

- Step 1** Run the `--jars` command to load the `flume-ng-sdk-{version}.jar` dependency package.
  - Step 2** Modify the two configuration items in the `spark-default.conf` file:  
`spark.driver.extraClassPath=$PWD/*: {Add the original value}`  
`spark.executor.extraClassPath = $PWD/*`
  - Step 3** Run the job successfully. If an error is reported, check which JAR is not loaded and perform step 1 and step 2 again.
- End

## 10.2 Failed to Install a Flume Client

### Symptom

A Flume client fails to be installed, and "JAVA\_HOME is null" or "flume has been installed" is displayed.

```
CST 2016-08-31 17:02:51 [flume-client install]: JAVA_HOME is null in current user,please install the JDK and set the JAVA_HOME
```

```
CST 2016-08-31 17:02:51 [flume-client install]: check environment failed.
```

```
CST 2016-08-31 17:02:51 [flume-client install]: check param failed.
```

```
CST 2016-08-31 17:02:51 [flume-client install]: install flume client failed.
```

```
CST 2016-08-31 17:03:58 [flume-client install]: flume has been installed
```

```
CST 2016-08-31 17:03:58 [flume-client install]: check path failed.
```

```
CST 2016-08-31 17:03:58 [flume-client install]: check param failed.
```

```
CST 2016-08-31 17:03:58 [flume-client install]: install flume client failed.
```

### Cause Analysis

- Environment variables are checked during Flume client installation. If no Java is available, an error message is displayed and the installation quits.
- The Flume client has been installed in the specified directory.

### Solution

**Step 1** Run the following command if an error message is displayed stating "JAVA\_HOME is null":

```
export JAVA_HOME=Java path
```

Set **JAVA\_HOME** and execute the installation script again.

**Step 2** If a Flume client has been installed under the specified directory, uninstall the client and use another directory.

----End

## 10.3 A Flume Client Cannot Connect to the Server

### Symptom

A user installs a Flume client and sets an Avro sink to communicate with the server. However, the Flume server cannot be connected.

### Cause Analysis

1. The server is incorrectly configured and the monitoring port fails to be started. For example, an incorrect IP address or an occupied port is configured for the Avro source of the server.

View Flume run logs.

```
2016-08-31 17:28:42,092 | ERROR | [lifecycleSupervisor-1-9] | Unable to start EventDrivenSourceRunner: { source:Avro source avro_source: { bindAddress: 10.120.205.7, port: 21154 } } - Exception follows. | org.apache.flume.lifecycle.LifecycleSupervisor
```

```
$MonitorRunnable.run(LifecycleSupervisor.java:253)  
java.lang.RuntimeException: org.jboss.netty.channel.ChannelException: Failed to bind to: /  
192.168.205.7:21154
```

2. If encrypted transmission is used, the certificate or password is incorrect.  
2016-08-31 17:15:59,593 | ERROR | [conf-file-poller-0] | Source avro\_source has been removed due to an error during configuration |  
org.apache.flume.node.AbstractConfigurationProvider.loadSources(AbstractConfigurationProvider.java:388)  
org.apache.flume.FlumeException: Avro source configured with invalid keystore: /opt/Bigdata/MRS\_XXX/install/FusionInsight-Flume-1.9.0/flume/conf/flume\_sChat.jks
3. The network connection between the client and the server is abnormal.  
PING 192.168.85.55 (10.120.85.55) 56(84) bytes of data.  
From 192.168.85.50 icmp\_seq=1 Destination Host Unreachable  
From 192.168.85.50 icmp\_seq=2 Destination Host Unreachable  
From 192.168.85.50 icmp\_seq=3 Destination Host Unreachable  
From 192.168.85.50 icmp\_seq=4 Destination Host Unreachable

## Solution

- Step 1** Set a correct IP address (an IP address of the local host). If the port has been occupied, configure another free port.
- Step 2** Configure a correct certificate path.
- Step 3** Contact the network administrator to restore the network.

----End

## 10.4 Flume Data Fails to Be Written to the Component

### Symptom

After the Flume process is started, Flume data cannot be written to the target component (in this example, data is to be written to HDFS).

### Cause Analysis

1. HDFS is not started or is faulty.

View Flume run logs.

```
2019-02-26 11:16:33,564 | ERROR | [SinkRunner-PollingRunner-DefaultSinkProcessor] | operation the  
hdfs file errors. | org.apache.flume.sink.hdfs.HDFSEventSink.process(HDFSEventSink.java:414)  
2019-02-26 11:16:33,747 | WARN | [hdfs-CCCC-call-runner-4] | A failover has occurred since the start  
of call #32795 ClientNamenodeProtocolTranslatorPB.getFileInfo over  
192-168-13-88/192.168.13.88:25000 | org.apache.hadoop.io.retry.RetryInvocationHandler  
$ProxyDescriptor.failover(RetryInvocationHandler.java:220)  
2019-02-26 11:16:33,748 | ERROR | [hdfs-CCCC-call-runner-4] | execute hdfs error. {} |  
org.apache.flume.sink.hdfs.HDFSEventSink$3.call(HDFSEventSink.java:744)  
java.net.ConnectException: Call From 192-168-12-221/192.168.12.221 to 192-168-13-88:25000 failed  
on connection exception: java.net.ConnectException: Connection refused; For more details see: http://  
wiki.apache.org/hadoop/ConnectionRefused
```

2. The HDFS sink is not started.

Check the Flume run log. It is found that **Flume current metrics** does not contain sink information.

```
2019-02-26 11:46:05,501 | INFO | [pool-22-thread-1] | flume current metrics:{"CHANNEL.BBBB":  
{"ChannelCapacity":10000,"ChannelFillPercentage":0.0,"Type":"CHANNEL","ChannelStoreSize":0,"  
EventProcessTimedelta":0,"EventTakeSuccessCount":0,"ChannelSize":0,"EventTakeAttemptCount":  
0,"StartTime":1551152734999,"EventPutAttemptCount":0,"EventPutSuccessCount":0,"StopTime":  
0},"SOURCE.AAAA":  
{"AppendBatchAcceptedCount":0,"EventAcceptedCount":0,"AppendReceivedCount":0,"MonTime":
```

```
"0","StartTime":"1551152735503","AppendBatchReceivedCount":"0","EventReceivedCount":"0","Type":"SOURCE","TotalFilesCount":"1001","SizeAcceptedCount":"0","UpdateTime":"605410241202740","AppendAcceptedCount":"0","OpenConnectionCount":"0","MovedFilesCount":"1001","StopTime":"0"} | org.apache.flume.node.Application.getRestartComps(Application.java:467)
```

## Solution

- Step 1** If the component to which Flume writes data is not started, start the component. If the component is abnormal, contact technical support.
- Step 2** If the sink is not started, check whether the configuration file is correctly configured. If it is not, correct the configuration file and restart the Flume process. If it is correctly configured, view the error information in the log and rectify the fault based on the error information.

----End

## 10.5 Flume Server Process Fault

### Symptom

After Flume runs for a period of time, the Flume instance is in the faulty state on Manager.

### Cause Analysis

If the Flume file or folder permission is abnormal, the following information is displayed on MRS Manager after the restart:

```
[2019-02-26 13:38:02]RoleInstance prepare to start failure [{ScriptExecutionResult=ScriptExecutionResult [exitCode=126, output=, errMsg=sh: line 1: /opt/Bigdata/MRS_XXX/install/FusionInsight-Flume-1.9.0/flume/bin/flume-manage.sh: Permission denied
```

### Solution

Compare the file and folder permissions with those for the Flume node that is running properly and correct the file or folder permissions.

## 10.6 Flume Data Collection Is Slow

### Symptom

After Flume is started, it takes a long time for Flume to collect data.

### Cause Analysis

1. The heap memory of Flume is not properly set. As a result, the Flume process keeps in the GC state. View Flume run logs.  
2019-02-26T13:06:20.666+0800: 1085673.512: [Full GC:[CMS: 3849339k->3843458K(3853568K), 2.5817610 secs] 4153654K->3843458K(4160256K), [CMS Perm : 27335K->27335K(45592K),2.5820080 SECS] [Times: user=2.63, sys0.00, real=2.59 secs]
2. The **deletePolicy** policy configured for the Spooldir source is **immediate**.

## Solution

**Step 1** Increase the size of the heap memory (**xmx**).

**Step 2** Change the **deletePolicy** policy of the SpoolDir source to **never**.

----End

## 10.7 Failed to Start Flume

### Symptom

The Flume service fails to be installed or restarted.

### Cause Analysis

1. The heap memory of Flume is greater than the remaining memory of the server. The Flume startup log shows the following information:  
[CST 2019-02-26 13:31:43][INFO] [[checkMemoryValidity:124]] [GC\_OPTS is invalid: Xmx(40960000MB) is bigger than the free memory(56118MB) in system.] [9928]
2. The permission on the Flume file or folder is abnormal. The following information is displayed on the GUI or in the background:  
[2019-02-26 13:38:02]RoleInstance prepare to start failure  
[[ScriptExecutionResult=ScriptExecutionResult [exitCode=126, output=, errMsg=sh: line 1: /opt/Bigdata/MRS\_XXX/install/FusionInsight-Flume-1.9.0/flume/bin/flume-manage.sh: Permission denied
3. The **JAVA\_HOME** is incorrectly configured. The Flume agent startup log shows the following information:  
Info: Sourcing environment configuration script /opt/FlumeClient/fusioninsight-flume-1.9.0/conf/flume-env.sh  
+ '[' -n '' ]'  
+ exec /tmp/MRS-Client/MRS\_Flume\_ClientConfig/JDK/jdk-8u18/bin/java '-  
XX:OnOutOfMemoryError=bash /opt/FlumeClient/fusioninsight-flume-1.9.0/bin/  
out\_memory\_error.sh /opt/FlumeClient/fusioninsight-flume-1.9.0/conf %p' -Xms2G -Xmx4G -  
XX:CMSFullGCsBeforeCompaction=1 -XX:+UseConcMarkSweepGC -XX:+CMSParallelRemarkEnabled -  
XX:+UseCMSCompactAtFullCollection -Dkerberos.domain.name=hadoop.hadoop.com -verbose:gc -  
XX:+UseGCLogFileRotation -XX:NumberOfGCLogFiles=10 -XX:GCLogFileSize=1M -XX:+PrintGCDetails -  
XX:+PrintGCDateStamps -Xloggc:/var/log/Bigdata//flume-client-1/flume/flume-root-20190226134231-  
%p-gc.log -Dproc\_org.apache.flume.node.Application -Dproc\_name=client -Dproc\_conf\_file=/opt/  
FlumeClient/fusioninsight-flume-1.9.0/conf/properties.properties -Djava.security.krb5.conf=/opt/  
FlumeClient/fusioninsight-flume-1.9.0/conf/krb5.conf -Djava.security.auth.login.config=/opt/  
FlumeClient/fusioninsight-flume-1.9.0/conf/jaas.conf -Dzookeeper.server.principal=zookeeper/  
hadoop.hadoop.com -Dzookeeper.request.timeout=120000 -Dflume.instance.id=884174180 -  
Dflume.agent.name=clientName1 -Dflume.role=client -Dlog4j.configuration.watch=true -  
Dlog4j.configuration=log4j.properties -Dflume\_log\_dir=/var/log/Bigdata//flume-client-1/flume/ -  
Dflume.service.id=flume-client-1 -Dbeetle.application.home.path=/opt/FlumeClient/fusioninsight-  
flume-1.9.0/conf/service -Dflume.called.from.service -Dflume.conf.dir=/opt/FlumeClient/fusioninsight-  
flume-1.9.0/conf -Dflume.metric.conf.dir=/opt/FlumeClient/fusioninsight-flume-1.9.0/conf -  
Dflume.script.home=/opt/FlumeClient/fusioninsight-flume-1.9.0/bin -cp '/opt/FlumeClient/  
fusioninsight-flume-1.9.0/conf:/opt/FlumeClient/fusioninsight-flume-1.9.0/lib/\*:/opt/FlumeClient/  
fusioninsight-flume-1.9.0/conf/service/' -Djava.library.path=/opt/FlumeClient/fusioninsight-flume-1.9.0/  
plugins.d/native/native org.apache.flume.node.Application --conf-file /opt/FlumeClient/fusioninsight-  
flume-1.9.0/conf/properties.properties --name client  
/opt/FlumeClient/fusioninsight-flume-1.9.0/bin/flume-ng: line 233: /tmp/FusionInsight-Client/Flume/  
FusionInsight\_Flume\_ClientConfig/JDK/jdk-8u18/bin/java: No such file or directory

### Solution

**Step 1** Increase the size of the heap memory (**xmx**).

**Step 2** Compare the file and folder permissions with those for node where Flume is started properly and change the incorrect file or folder permissions.

**Step 3** Reconfigure **JAVA\_HOME**.

On the client, replace the value of **JAVA\_HOME** in the **`\${install\_home}/fusioninsight-flume-*Flume version*/conf/ENV\_VARS** file. On the server, replace the value of **JAVA\_HOME** in the **ENV\_VARS** file in the **etc** directory.

To obtain the value of **JAVA\_HOME**, log in to the node where Flume is properly started and run the **echo `\${JAVA\_HOME}** command.

 **NOTE**

**`\${install\_home}** is the installation path of the Flume client.

----End

# 11 Using HBase

## 11.1 Slow Response to HBase Connection

### Symptom

In the same VPC, response is slow when an external cluster connects to HBase through Phoenix.

```
[root@node-master2-ko2b.j bin]# ./sqlite.py 192.168.1.109:2101
Setting property: {incremental, false}
Setting property: {isolation, TRANSACTION_READ_COMMITTED}
Issuing: 'connect jdbc:phoenix:192.168.1.109:2101 none none org.apache.phoenix.jdbc.PhoenixDriver'
Connecting to jdbc:phoenix:192.168.1.109:2101
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/apache-phoenix-4.13.0-HBase-1.3-bin/phoenix-4.13.0-HBase-1.3-client.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/share/slf4j-log4j12-1.7.10/slf4j-log4j12-1.7.10.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
19:01:17.17:29:34 INFO util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java class
es where applicable
Connected to: Phoenix (version 4.13)
Driver: PhoenixEmbeddedDriver (version 4.13)
Autocommit status: true
Transaction isolation: TRANSACTION_READ_COMMITTED
Building list of tables and columns for tab-completion (set fastconnect to true to skip)...
263/269 (100%) Done
Done
sqlite version 1.2.0
0: jdbc:phoenix:192.168.1.109:2101>
```

### Possible Cause

DNS has been configured. When a client connects to HBase, DNS resolves the server first, causing slow response.

### Procedure

- Step 1** Log in to the master node as user **root**.
- Step 2** Run the **vi /etc/resolv.conf** command to open the **resolv.conf** file and comment out the address of the DNS server.

For example, #1.1.1.1.

----End

## 11.2 Failed to Authenticate the HBase User

### Issue

Failed to authenticate the HBase user.

### Symptom

Failed to authenticate the HBase user on the client. The following error information is displayed:

```
2019-05-13 10:53:09,975 ERROR [localhost-startStop-1] xxxConfig.LoginUtil: login failed with hbaseuser and /usr/local/linoseyc/hbase-tomcat/webapps/bigdata_hbase/WEB-INF/classes/user.keytab.
2019-05-13 10:53:09,975 ERROR [localhost-startStop-1] xxxConfig.LoginUtil: perhaps cause 1 is (wrong password) keytab file and user not match, you can kinit -k -t keytab user in client server to check.
2019-05-13 10:53:09,975 ERROR [localhost-startStop-1] xxxConfig.LoginUtil: perhaps cause 2 is (clock skew) time of local server and remote server not match, please check ntp to remote server.
2019-05-13 10:53:09,975 ERROR [localhost-startStop-1] xxxConfig.LoginUtil: perhaps cause 3 is (aes256 not support) aes256 not support by default jdk/jre, need copy local_policy.jar and US_export_policy.jar from remote server in path ${BIGDATA_HOME}/jdk/jre/lib/security.
```

### Cause Analysis

The version of the JAR file in the JDK used by the user is different from that of the JAR file authenticated by MRS.

### Procedure

- Step 1** Log in to the Master1 node as user **root**.
- Step 2** Run the following command to check the JAR file authenticated by MRS:  

```
ll /opt/share/local_policy/local_policy.jar
```

```
ll /opt/Bigdata/jdk{version}/jre/lib/security/local_policy.jar
```
- Step 3** Download the JAR package queried in step 2 to the local host.
- Step 4** Copy the downloaded JAR package to the local JDK directory **/opt/Bigdata/jdk/jre/lib/security**.
- Step 5** Run the **cd Client installation directory/HBase/hbase/bin** command to go to the **bin** directory of HBase.
- Step 6** Run the **sh start-hbase.sh** command to restart HBase.  
----End

## 11.3 RegionServer Failed to Start Because the Port Is Occupied

### Symptom

RegionServer is in the **Restoring** state on Manager.



## Cause Analysis

1. View the RegionServer log (`/var/log/Bigdata/hbase/rs/hbase-omm-xxx.log`).
2. Run the `lsof -i:21302` command (the port number of MRS 1.7.X and later versions is 16020) to view the PID. Based on the PID, check the process. It is found that the RegionServer port is occupied by DFSZkFailoverController.
3. The value of `/proc/sys/net/ipv4/ip_local_port_range` is **9000 65500**. The temporary port range and the MRS port range overlap. This is because the preinstall operation is not performed during installation.

## Solution

- Step 1** Run the `kill -9 DFSZkFailoverController pid` command to ensure that another port is bound with after a restart and restart the RegionServer in the **Restoring** state.

----End

# 11.4 HBase Failed to Start Due to Insufficient Node Memory

## Symptom

The RegionServer service of HBase is always in the **Restoring** state.

## Cause Analysis

1. Check the RegionServer log (`/var/log/Bigdata/hbase/rs/hbase-omm-XXX.out`). It is found that the following information is printed:  
There is insufficient memory for the Java Runtime Environment to continue.
2. Run the `free` command to check the memory. It is found that the available memory of the node is insufficient.

## Solution

- Step 1** Locate why the memory is insufficient. It is found that some processes occupy too much memory or the server does not have sufficient memory.

----End

# 11.5 HBase Service Unavailable Due to Poor HDFS Performance

## Symptom

The HBase component continuously reports alarms indicating that the service is unavailable.

## Cause Analysis

HDFS performance is low, causing health check timeout and the alarm is generated accordingly. You can perform the following operations:

1. View the HMaster log (`/var/log/Bigdata/hbase/hm/hbase-omm-xxx.log`) and check that **system pause**, **jvm**, and other GC-related information is not frequently printed in the log.
2. Determine whether the fault is caused by poor HDFS performance using either of the following methods:
  - a. Run **hbase shell** to access the HBase shell, and run the **list** command to check whether it takes a long period of time to list all tables in HBase.
  - b. Enable printing of the debug logs of HDFS, and check whether it takes a long period of time to list the content of a large number of directories by running the **hadoop fs -ls /XXX/XXX** command.
  - c. Print the Java stack information about a specified HMaster process.  
**su - omm**  
**jps**  
**jstack pid**
3. Check the jstack information. The following figure shows that the process is stuck at the **DFSClient.listPaths** state.

Figure 11-1 Exception

```
java.lang.Thread.State: WAITING (on object monitor)
  at java.lang.Object.wait(Native Method)
  at java.lang.Object.wait(Object.java:503)
  at org.apache.hadoop.ipc.Client.call(Client.java:1396)
  - locked <0x0000000b9268a38> (a org.apache.hadoop.ipc.Client$Call)
  at org.apache.hadoop.ipc.Client.call(Client.java:1363)
  at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:206)
  at com.sun.proxy.$Proxy13.getListing(Unknown Source)
  at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.getListing(ClientNamenodeProtocolTranslatorPB.java:187)
  at sun.reflect.GeneratedMethodAccessor24.invoke(Unknown Source)
  at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
  at java.lang.reflect.Method.invoke(Method.java:606)
  at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:187)
  at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:102)
  at com.sun.proxy.$Proxy14.getListing(Unknown Source)
  at sun.reflect.GeneratedMethodAccessor24.invoke(Unknown Source)
  at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
  at java.lang.reflect.Method.invoke(Method.java:606)
  at org.apache.hadoop.hbase.fs.HFileSystem$1.invoke(HFileSystem.java:294)
  at com.sun.proxy.$Proxy17.getListing(Unknown Source)
  at org.apache.hadoop.hdfs.DFSClient.listPaths(DFSClient.java:1767)
  at org.apache.hadoop.hdfs.DFSClient.listPaths(DFSClient.java:1750)
  at org.apache.hadoop.hdfs.DistributedFileSystem.listStatusInternal(DistributedFileSystem.java:691)
  at org.apache.hadoop.hdfs.DistributedFileSystem.access$600(DistributedFileSystem.java:102)
  at org.apache.hadoop.hdfs.DistributedFileSystem$15.doCall(DistributedFileSystem.java:753)
  at org.apache.hadoop.hdfs.DistributedFileSystem$15.doCall(DistributedFileSystem.java:749)
  at org.apache.hadoop.fs.FileSystemLinkResolver.resolve(FileSystemLinkResolver.java:81)
  at org.apache.hadoop.hdfs.DistributedFileSystem.listStatus(DistributedFileSystem.java:749)
  at org.apache.hadoop.fs.FileSystem.listStatus(FileSystem.java:1483)
```

## Solution

**Step 1** If this alarm is caused by poor HDFS performance, check whether Impala is of an earlier version or JournalNode was incorrectly deployed during the initial deployment (more than three JournalNode nodes are deployed).

----End



## Solution

- Step 1** Log in to the node and run the `ps -ef | grep HRegionServer` command in the background. A residual process is found.
- Step 2** After confirming that the process can be killed, kill the process. If the process cannot be stopped by running the `kill` command, run the `kill -9` command to forcibly stop the process.
- Step 3** Restart the HBase service.

----End

# 11.8 HBase Failed to Start Due to a Quota Set on HDFS

## Symptom

HBase fails to start.

## Cause Analysis

Check the HMaster log (`/var/log/Bigdata/hbase/hm/hbase-omm-xxx.log`). It is found that "The DiskSpace quota of /hbase is exceeded" is displayed.

```
Caused by:
org.apache.hadoop.hdfs.protocol.DSQuotaExceededException: The DiskSpace quota of /hbase is exceeded: quota=29240.3g diskSpace consumed=37945.7g
at org.apache.hadoop.hdfs.server.namenode.HNodeDirectoryWithQuota.verifyQuota(HNodeDirectoryWithQuota.java:159)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.verifyQuota(FSDirectory.java:1643)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.updateCount(FSDirectory.java:1378)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.addChild(FSDirectory.java:1745)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.addChild(FSDirectory.java:1762)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.unprotectedMkdir(FSDirectory.java:1581)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.mkdirs(FSDirectory.java:1537)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.mkdirsInternal(FSNamesystem.java:2768)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.mkdirs(FSNamesystem.java:2724)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.mkdirs(NameNodeRpcServer.java:641)
at org.apache.hadoop.hdfs.protocolPB.ClientNameNodeProtocolServerSideTranslatorPB.mkdirs(ClientNameNodeProtocolServerSideTranslatorPB.java:416)
at org.apache.hadoop.hdfs.protocol.proto.ClientNameNodeProtocolProtosClientNameNodeProtocol$2.callBlockingMethod(ClientNameNodeProtocolProtos$2)
at org.apache.hadoop.hdfs.protocol.proto.ClientNameNodeProtocolProtosClientNameNodeProtocol$2.call(ProtocolEngine.java:427)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:427)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:925)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:1740)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:1706)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:415)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1332)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:1704)

at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:57)
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:525)
at org.apache.hadoop.ipc.RemoteException.instantiateException(RemoteException.java:90)
at org.apache.hadoop.ipc.RemoteException.unwrapRemoteException(RemoteException.java:57)
at org.apache.hadoop.hdfs.DFSClient.mkdirs(DFSClient.java:1858)
at org.apache.hadoop.hdfs.DistributedFileSystem.mkdirs(DistributedFileSystem.java:483)
at org.apache.hadoop.fs.FileSystem.mkdirs(FileSystem.java:1728)
at org.apache.hadoop.hbase.RegionServer.wal.HLog.<init>(HLog.java:413)
at org.apache.hadoop.hbase.RegionServer.wal.HLog.<init>(HLog.java:397)
at org.apache.hadoop.hbase.regionserver.HRegionServer.initializeHLog(HRegionServer.java:1348)
at org.apache.hadoop.hbase.regionserver.HRegionServer.setUpAllAndReplication(HRegionServer.java:1337)
at org.apache.hadoop.hbase.regionserver.HRegionServer.handleReportForDutyResponse(HRegionServer.java:1048)
at org.apache.hadoop.hbase.regionserver.HRegionServer.run(HRegionServer.java:714)
at java.lang.Thread.run(Thread.java:722)
```

## Solution

- Step 1** Run the `df -h` command to check data directory space. It is found that the directory space is full. Delete unnecessary data to free up space.
- Step 2** Expand the node to ensure that the data directory space is sufficient.

----End

## 11.9 HBase Failed to Start Due to Corrupted Version Files

### Symptom

HBase fails to start.

### Cause Analysis

1. The **hbase.version** file is read during HBase startup. However, the log indicates that a reading exception occurs.

```
2019-07-27 15:38:18,692 | ERROR | master/node-master1r2w:16080:becomeActiveMaster | Failed to become active master | org.slf4j.helpers.MarkerIgnoringBase.error(MarkerIgnoringBase.java:159)
org.apache.hadoop.hbase.util.FileSystemVersionException: HBase file layout needs to be upgraded. You have version null and I want version 8. Consult http://hbase.apache.org/book.html for further information about upgrading HBase. Is your hbase.rootdir valid? If so, you may need to run 'hbase hbck -fixVersionFile'.
    at org.apache.hadoop.hbase.util.FSUtils.checkVersion(FSUtils.java:509)
    at org.apache.hadoop.hbase.master.MasterFileSystem.checkRootDir(MasterFileSystem.java:271)
    at org.apache.hadoop.hbase.master.MasterFileSystem.createInitialFileSystemLayout(MasterFileSystem.java:151)
    at org.apache.hadoop.hbase.master.MasterFileSystem.<init>(MasterFileSystem.java:122)
    at org.apache.hadoop.hbase.master.HMaster.finishActiveMasterInitialization(HMaster.java:869)
    at org.apache.hadoop.hbase.master.HMaster.startActiveMasterManager(HMaster.java:2207)
```

2. The file cannot be viewed by running the **hadoop fs -cat /hbase/hbase.version** command. The file is corrupted.

### Solution

**Step 1** Run the **hbase hbck -fixVersionFile** command to restore the file.

**Step 2** If the problem persists after performing **Step 1**, obtain the **hbase.version** file from another cluster of the same version and upload the file to replace the original one.

**Step 3** Restart the HBase service.

----End

## 11.10 High CPU Usage Caused by Zero-Loaded RegionServer

### Symptom

The CPU usage of RegionServer is high, but there is no service running on RegionServer.

### Cause Analysis

1. Run the **top** command to obtain the CPU usage of RegionServer processes and check the IDs of processes with high CPU usage.
2. Obtain the CPU usage of threads under these processes based on the RegionServer process IDs.

Run the **top -H -p <PID>** (replace it with the actual RegionServer process ID). As shown in the following figure, the CPU usage of some threads reaches 80%.

| PID   | USER | PR | NI | VIRT    | RES  | SHR   | S | %CPU | %MEM | TIME+   | COMMAND |
|-------|------|----|----|---------|------|-------|---|------|------|---------|---------|
| 75706 | omm  | 20 | 0  | 6879444 | 1.0g | 25612 | S | 90.4 | 1.6  | 0:00.00 | java    |
| 75716 | omm  | 20 | 0  | 6879444 | 1.0g | 25612 | S | 90.4 | 1.6  | 0:04.74 | java    |
| 75720 | omm  | 20 | 0  | 6879444 | 1.0g | 25612 | S | 88.6 | 1.6  | 0:01.93 | java    |

```
75721 omm 20 0 6879444 1.0g 25612 S 86.8 1.6 0:01.99 java
75722 omm 20 0 6879444 1.0g 25612 S 86.8 1.6 0:01.94 java
75723 omm 20 0 6879444 1.0g 25612 S 86.8 1.6 0:01.96 java
75724 omm 20 0 6879444 1.0g 25612 S 86.8 1.6 0:01.97 java
75725 omm 20 0 6879444 1.0g 25612 S 81.5 1.6 0:02.06 java
75726 omm 20 0 6879444 1.0g 25612 S 79.7 1.6 0:02.01 java
75727 omm 20 0 6879444 1.0g 25612 S 79.7 1.6 0:01.95 java
75728 omm 20 0 6879444 1.0g 25612 S 78.0 1.6 0:01.99 java
```

- Obtain the thread stack information based on the ID of the RegionServer process.

**jstack 12345 >allstack.txt** (Replace it with the actual RegionServer process ID.)

- Convert the thread ID into the hexadecimal format:

**printf "%x\n" 30648**

In the command output, the TID is **77b8**.

- Search the thread stack based on the hexadecimal TID. It is found that the compaction operation is performed.

```
"regionserver/ahbd-hbase-dat1/12.2.168.1:21302-longCompactions-1482676601478" #1641 prio=5 os_prio=0 tid=0x00007fa614563000 nid=0x77b8 runnable [0x0]
java.lang.Thread.State: RUNNABLE
  at org.apache.hadoop.io.compress.snappy.SnappyCompressor.compressByteDirect(Native Method)
  at org.apache.hadoop.io.compress.snappy.SnappyCompressor.compress(SnappyCompressor.java:228)
  at org.apache.hadoop.io.compress.BlockCompressorStream.compress(BlockCompressorStream.java:149)
  at org.apache.hadoop.io.compress.BlockCompressorStream.finish(BlockCompressorStream.java:142)
  at org.apache.hadoop.hbase.io.encoding.HFileBlockDefaultEncodingContext.compressAfterEncoding(HFileBlockDefaultEncodingContext.java:219)
  at org.apache.hadoop.hbase.io.encoding.HFileBlockDefaultEncodingContext.compressAndEncrypt(HFileBlockDefaultEncodingContext.java:132)
  at org.apache.hadoop.hbase.io.hfile.HFileBlock$Writer.finishBlock(HFileBlock.java:989)
  at org.apache.hadoop.hbase.io.hfile.HFileBlock$Writer.ensureBlockReady(HFileBlock.java:961)
  at org.apache.hadoop.hbase.io.hfile.HFileBlock$Writer.finishBlockAndWriteHeaderAndData(HFileBlock.java:1077)
```

- Perform the same operations on other threads. It is found that the threads are compaction threads.

```
"regionserver/ahbd-hbase-dat1/12.2.168.1:21302-longCompactions-1482676601473" #1629 prio=5 os_prio=0 tid=0x00007fa61454d800 nid=0x77a0 runnable
java.lang.Thread.State: RUNNABLE
  at org.apache.hadoop.hdfs.DFSOutputStream.writeChunk(DFSOutputStream.java:425)
  - locked <0x000000020276ba38> (a org.apache.hadoop.hdfs.DFSOutputStream)
  at org.apache.hadoop.fs.FSOutputSummar.writeChecksumChunks(FSOutputSummar.java:214)
  at org.apache.hadoop.fs.FSOutputSummar.flushBuffer(FSOutputSummar.java:165)
  - locked <0x000000020276ba38> (a org.apache.hadoop.hdfs.DFSOutputStream)
  at org.apache.hadoop.fs.FSOutputSummar.flushBuffer(FSOutputSummar.java:146)
  - eliminated <0x000000020276ba38> (a org.apache.hadoop.hdfs.DFSOutputStream)
  at org.apache.hadoop.fs.FSOutputSummar.write1(FSOutputSummar.java:137)
  at org.apache.hadoop.fs.FSOutputSummar.write(FSOutputSummar.java:112)
  - locked <0x000000020276ba38> (a org.apache.hadoop.hdfs.DFSOutputStream)
  at org.apache.hadoop.fs.FSDataOutputStream$PositionCache.write(FSDataOutputStream.java:58)
  at java.io.DataOutputStream.write(DataOutputStream.java:107)
  - locked <0x00000004de9535c8> (a org.apache.hadoop.hdfs.client.HdfsDataOutputStream)
  at java.io.FilterOutputStream.write(FilterOutputStream.java:97)
```

## Solution

This is a normal phenomenon.

The threads that consume a large number of CPU resources are compaction threads. Some threads invoke the Snappy compression algorithm, and some threads invoke HDFS data writing and reading. Each region has massive sets of data and numerous data files and uses the Snappy compression algorithm. For this reason, the compaction operations consume a large number of CPU resources.

## Fault Locating Methods

**Step 1** Run the **top** command to check the process with high CPU usage.

**Step 2** Check the threads with high CPU usage in the process.

Run the **top -H -p <PID>** command to print CPU usage of threads under the process.

Obtain the thread with the highest CPU usage from the query result. You can also obtain the thread by running the following command:

Or run the **ps -mp <PID> -o THREAD,tid,time | sort -rn** command.

View the command output to obtain the ID of the thread with the highest CPU usage.

**Step 3** Obtain the stack of the faulty thread.

The **jstack** tool is the most effective and reliable tool for locating Java problems.

You can obtain the **jstack** tool from the **java/bin** directory.

```
jstack <PID> > allstack.txt
```

Obtain the process stack and output it to a local file.

**Step 4** Convert the thread ID into the hexadecimal format:

```
printf "%x\n" <PID>
```

The process ID in the command output is the TID.

**Step 5** Run the following command to obtain the TID and output it to a local file:

```
jstack <PID> | grep <TID> > Onestack.txt
```

If you want to view the TID in the CLI only, run the following command:

```
jstack <PID> | grep <TID> -A 30
```

-A 30 indicates that 30 lines are displayed.

----End

## 11.11 HBase Failed to Start with "FileNotFoundException" in RegionServer Logs

### Symptom

HBase fails to start, and the RegionServer stays in the **Restoring** state.

### Cause Analysis

1. Check the RegionServer log (**/var/log/Bigdata/hbase/rs/hbase-omm-XXX.out**). It is found that the following information is printed:

```
| ERROR | RS_OPEN_REGION-ab-dn01:21302-2 | ABORTING region server ab-  
dn01,21302,1487663269375: The coprocessor  
org.apache.kylin.storage.hbase.cube.v2.coprocessor.endpoint.CubeVisitService threw  
java.io.FileNotFoundException: File does not exist: hdfs://hacluster/kylin/kylin_metadata/coprocessor/  
kylin-coprocessor-1.6.0-SNAPSHOT-0.jar |  
org.apache.hadoop.hbase.regionserver.HRegionServer.abort(HRegionServer.java:2123)  
java.io.FileNotFoundException: File does not exist: hdfs://hacluster/kylin/kylin_metadata/coprocessor/  
kylin-coprocessor-1.6.0-SNAPSHOT-0.jar  
at org.apache.hadoop.hdfs.DistributedFileSystem$25.doCall(DistributedFileSystem.java:1399)  
at org.apache.hadoop.hdfs.DistributedFileSystem$25.doCall(DistributedFileSystem.java:1391)  
at org.apache.hadoop.fs.FileSystemLinkResolver.resolve(FileSystemLinkResolver.java:81)  
at org.apache.hadoop.hdfs.DistributedFileSystem.getFileStatus(DistributedFileSystem.java:1391)  
at org.apache.hadoop.fs.FileUtil.copy(FileUtil.java:340)  
at org.apache.hadoop.fs.FileUtil.copy(FileUtil.java:292)  
at org.apache.hadoop.fs.FileSystem.copyToLocalFile(FileSystem.java:2038)  
at org.apache.hadoop.fs.FileSystem.copyToLocalFile(FileSystem.java:2007)  
at org.apache.hadoop.fs.FileSystem.copyToLocalFile(FileSystem.java:1983)
```

```
at org.apache.hadoop.hbase.util.CoprocessorClassLoader.init(CoprocessorClassLoader.java:168)
at
org.apache.hadoop.hbase.util.CoprocessorClassLoader.getClassLoader(CoprocessorClassLoader.java:250)
at org.apache.hadoop.hbase.coprocessor.CoprocessorHost.load(CoprocessorHost.java:224)
at
org.apache.hadoop.hbase.regionserver.RegionCoprocessorHost.loadTableCoprocessors(RegionCoprocessorHost.java:365)
at
org.apache.hadoop.hbase.regionserver.RegionCoprocessorHost.<init>(RegionCoprocessorHost.java:227)
at org.apache.hadoop.hbase.regionserver.HRegion.<init>(HRegion.java:783)
at org.apache.hadoop.hbase.regionserver.HRegion.<init>(HRegion.java:689)
at sun.reflect.GeneratedConstructorAccessor22.newInstance(Unknown Source)
at
sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:423)
at org.apache.hadoop.hbase.regionserver.HRegion.newHRegion(HRegion.java:6312)
at org.apache.hadoop.hbase.regionserver.HRegion.openHRegion(HRegion.java:6622)
at org.apache.hadoop.hbase.regionserver.HRegion.openHRegion(HRegion.java:6594)
at org.apache.hadoop.hbase.regionserver.HRegion.openHRegion(HRegion.java:6550)
at org.apache.hadoop.hbase.regionserver.HRegion.openHRegion(HRegion.java:6501)
at
org.apache.hadoop.hbase.regionserver.handler.OpenRegionHandler.openRegion(OpenRegionHandler.java:363)
at
org.apache.hadoop.hbase.regionserver.handler.OpenRegionHandler.process(OpenRegionHandler.java:129)
at org.apache.hadoop.hbase.executor.EventHandler.run(EventHandler.java:129)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
at java.lang.Thread.run(Thread.java:745)
```

2. Run the **hdfs** command on the client. It is found that the following file does not exist:  
`hdfs://hacluster/kylin/kylin_metadata/coprocessor/kylin-coprocessor-1.6.0-SNAPSHOT-0.jar`
3. When configuring the coprocessor for HBase, make sure that the path of the corresponding JAR package is correct. Otherwise, HBase cannot be started.

## Solution

Use the Apache Kylin engine to interconnect with MRS and make sure that the JAR file of the Kylin engine exists.

## 11.12 The Number of RegionServers Displayed on the Native Page Is Greater Than the Actual Number After HBase Is Started

### Symptom

After HBase is started, the number of RegionServers displayed on the HMaster native page is greater than the actual number.

The HMaster native page shows that four RegionServers are online, as shown in the following figure.



| ServerName                                 | Start time                   | Requests Per Second | Num. Regions |
|--------------------------------------------|------------------------------|---------------------|--------------|
| controller-192-168-1-1,21302,1494933959261 | Tue May 16 19:25:59 CST 2017 | 0                   | 19           |
| controller-192-168-1-2,21302,1494933957536 | Tue May 16 19:25:57 CST 2017 | 0                   | 24           |
| controller-192-168-1-3,21302,1494933958592 | Tue May 16 19:25:58 CST 2017 | 0                   | 16           |
| eth0,21302,1494933958592                   | Tue May 16 19:25:58 CST 2017 | 0                   | 0            |
| Total:4                                    |                              | 0                   | 59           |

## Cause Analysis

As shown in the following figure, the hostname of the node in the third row is **controller-192-168-1-3** and that of the fourth row is **eth0**. The two carry the same information reported by RegionServer. Then, log in to the corresponding nodes to check the **/etc/hosts** file. It is found that the same IP address is configured for the two hostnames. For details, see the following figure:

```
# special IPv6 addresses
::1          localhost ipv6-localhost ipv6-loopback
fe00::0     ipv6-localnet
ff00::0     ipv6-mcastprefix
ff02::1     ipv6-allnodes
ff02::2     ipv6-allrouters
ff02::3     ipv6-allhosts
11.1.1.3    eth2 eth2
#192.168.1.3 eth0 eth0
192.168.2.3 eth1 eth1
10.130.87.37 eth3 eth3
192.168.1.102 controller
1.1.1.1     hadoop.hadoop.com
192.168.1.2 controller-192-168-1-2
192.168.1.1 controller-192-168-1-1
192.168.1.3 controller-192-168-1-3
```

## Solution

Log in to the node where RegionServer resides, and modify the **/etc/hosts** file. Make sure that the same IP address can correspond to only one hostname.

## 11.13 RegionServer Instance Is in the Restoring State

### Symptom

HBase fails to start, and the RegionServer stays in the **Restoring** state.

## Cause Analysis

Check the running log (`/var/log/Bigdata/hbase/rs/hbase-omm-XXX.log`) of the abnormal RegionServer instance. It is found that the following information is displayed: **ClockOutOfSyncException..., Reported time is too far out of sync with master.**

```
2017-09-18 11:16:23,636 | FATAL | regionserver21302 | Master rejected startup because clock is out of sync |
org.apache.hadoop.hbase.regionserver.HRegionServer.reportForDuty(HRegionServer.java:2059)
org.apache.hadoop.hbase.ClockOutOfSyncException: org.apache.hadoop.hbase.ClockOutOfSyncException:
Server nl-bi-fi-datanode-24-65,21302,1505726180086 has been rejected; Reported time is too far out of
sync with master. Time difference of 152109ms > max allowed of 30000ms
at org.apache.hadoop.hbase.master.ServerManager.checkClockSkew(ServerManager.java:354)
...
...
2017-09-18 11:16:23,858 | ERROR | main | Region server exiting |
org.apache.hadoop.hbase.regionserver.HRegionServerCommandLine.start(HRegionServerCommandLine.java:
70)
java.lang.RuntimeException: HRegionServer Aborted
```

This log indicates that the time difference between the abnormal RegionServer instance and the HMaster instance is greater than the allowed time difference 30s (specified by the `hbase.regionserver.maxclockskew` parameter and the default value is **30000 ms**). As a result, the RegionServer instance is abnormal.

## Solution

Adjust the node time to ensure that the time difference between nodes is less than 30s.

# 11.14 HBase Failed to Start in a Newly Installed Cluster

## Symptom

HBase of a newly installed cluster fails to start. The RegionServer log contains the following error information:

```
2018-02-24 16:53:03,863 | ERROR | regionserver/host3/187.6.71.69:21302 | Master passed us a different
hostname to use; was=host3, but now=187-6-71-69 |
org.apache.hadoop.hbase.regionserver.HRegionServer.handleReportForDutyResponse(HRegionServer.java:138
6)
```

## Cause Analysis

In the `/etc/hosts` file, an IP address maps multiple hostnames.

## Solution

**Step 1** Modify the mapping between the IP address and hostnames in the `/etc/host` file.

**Step 2** Restart HBase.

----End

# 11.15 HBase Failed to Start Due to the Loss of the ACL Table Directory

## Symptom

The HBase cluster fails to start.

## Cause Analysis

1. Check the HMaster log of HBase. The following error information is displayed:

```

2018-04-10 09:14:05,610 | INFO | ftn-ies-301-a-f103:21300.activeMasterManager | Entered into preCreateTable. | org.apache.hadoop.hbase.index.coprocessor.mas
le(IndexMasterObserver.java:193)
2018-04-10 09:14:05,616 | INFO | ftn-ies-301-a-f103:21300.activeMasterManager | Exiting from preCreateTable. | org.apache.hadoop.hbase.index.coprocessor.mas
le(IndexMasterObserver.java:159)
2018-04-10 09:14:05,617 | INFO | ftn-ies-301-a-f103:21300.activeMasterManager | Client=null/null create 'hbase:acl', {NAME => 'l', BLOOMFILTER => 'NONE', VE
KEEP DELETED CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', CACHE_DATA_IN_L1 => 'true', MIN_VERSIONS => '0', BLOCK
, REPLICATION SCOPE => '0'} | org.apache.hadoop.hbase.master.HMaster.createTable(HMaster.java:1876)
2018-04-10 09:14:05,653 | ERROR | ftn-ies-301-a-f103:21300.activeMasterManager | Exception occurred while creating the table hbase:acl | org.apache.hadoop.hb
er.java:1909]
org.apache.hadoop.hbase.TableExistsException: hbase:acl
    at org.apache.hadoop.hbase.master.handler.CreateTableHandler.checkAndSetEnablingTable(CreateTableHandler.java:172)
    at org.apache.hadoop.hbase.master.handler.CreateTableHandler.prepare(CreateTableHandler.java:140)
    at org.apache.hadoop.hbase.master.HMaster.createTable(HMaster.java:1905)
    at org.apache.hadoop.hbase.security.access.AccessController.createACLTable(AccessController.java:128)
    at org.apache.hadoop.hbase.security.access.AccessController.postStartMaster(AccessController.java:1416)
    at org.apache.hadoop.hbase.master.MasterCoprocessorHost$62.call(MasterCoprocessorHost.java:769)
    at org.apache.hadoop.hbase.master.MasterCoprocessorHost.executeOperation(MasterCoprocessorHost.java:1315)
    at org.apache.hadoop.hbase.master.MasterCoprocessorHost.postStartMaster(MasterCoprocessorHost.java:765)
    at org.apache.hadoop.hbase.master.HMaster.finishActiveMasterInitialization(HMaster.java:933)
    at org.apache.hadoop.hbase.master.HMaster.access$900(HMaster.java:190)
    at org.apache.hadoop.hbase.master.HMaster$3.run(HMaster.java:2091)
    at java.lang.Thread.run(Thread.java:745)
2018-04-10 09:14:05,656 | ERROR | ftn-ies-301-a-f103:21300.activeMasterManager | Coprocessor postStartMaster() hook failed | org.apache.hadoop.hbase.master.H
ion(HMaster.java:935)
org.apache.hadoop.hbase.TableExistsException: hbase:acl
    at org.apache.hadoop.hbase.master.handler.CreateTableHandler.checkAndSetEnablingTable(CreateTableHandler.java:172)
    at org.apache.hadoop.hbase.master.handler.CreateTableHandler.prepare(CreateTableHandler.java:140)
    at org.apache.hadoop.hbase.master.HMaster.createTable(HMaster.java:1905)
    at org.apache.hadoop.hbase.security.access.AccessController.createACLTable(AccessController.java:128)
    at org.apache.hadoop.hbase.security.access.AccessController.postStartMaster(AccessController.java:1416)
    at org.apache.hadoop.hbase.master.MasterCoprocessorHost$62.call(MasterCoprocessorHost.java:769)
    at org.apache.hadoop.hbase.master.MasterCoprocessorHost.executeOperation(MasterCoprocessorHost.java:1315)

```

2. The HBase directory in HDFS is checked, which shows that the ACL table directory is lost.

### Browse Directory

| Permission | Owner | Group      | Size | Last Modified            | Replication | Block Size | Name      |
|------------|-------|------------|------|--------------------------|-------------|------------|-----------|
| drwx-----  | hbase | supergroup | 0 B  | Thu Mar 15 21:30:29 2018 | 0           | 0 B        | meta      |
| drwx-----  | hbase | supergroup | 0 B  | Thu Mar 15 21:30:36 2018 | 0           | 0 B        | namespace |

## Solution

**Step 1** Stop HBase.

**Step 2** Log in to the HBase client as the **hbase** user and run the following command.

```

Example:
hadoop03:~ # source /opt/client/bigdata_env
hadoop03:~ # kinit hbase
Password for hbase@HADOOP.COM:
hadoop03:~ # hbase zkcli

```

**Step 3** Delete the ACL table information from the ZooKeeper.

```

Example:
[zk: hadoop01:24002,hadoop02:24002,hadoop03:24002(CONNECTED) 0] deleteall /hbase/table/hbase:acl
[zk: hadoop01:24002,hadoop02:24002,hadoop03:24002(CONNECTED) 0] deleteall /hbase/table-lock/
hbase:acl

```

**Step 4** Start HBase.

----End

## 11.16 HBase Failed to Start After the Cluster Is Powered Off and On

### Symptom

After the ECS in the cluster is stopped and restarted, HBase fails to start.

### Cause Analysis

Check the HMaster run logs. A large number of errors are reported, as shown below:

```
2018-03-26 11:10:54,185 | INFO | hadoopc1h3,21300,1522031630949_splitLogManager__ChoreService_1 |
total tasks = 1 unassigned = 0 tasks={/hbase/splitWAL/WALs%2Fhadoopc1h1%2C213
02%2C1520214023667-splitting
%2Fhadoopc1h1%252C21302%252C1520214023667.default.1520584926990=last_update =
1522033841041 last_version = 34255 cur_worker_name = hadoopc1h3,21302,
1520943011826 status = in_progress incarnation = 3 resubmits = 3 batch = installed = 1 done = 0 error = 0}
| org.apache.hadoop.hbase.master.SplitLogManager$TimeoutMonitor.chore
(SplitLogManager.java:745)
2018-03-26 11:11:00,185 | INFO | hadoopc1h3,21300,1522031630949_splitLogManager__ChoreService_1 |
total tasks = 1 unassigned = 0 tasks={/hbase/splitWAL/WALs%2Fhadoopc1h1%2C213
02%2C1520214023667-splitting
%2Fhadoopc1h1%252C21302%252C1520214023667.default.1520584926990=last_update =
1522033841041 last_version = 34255 cur_worker_name = hadoopc1h3,21302,
1520943011826 status = in_progress incarnation = 3 resubmits = 3 batch = installed = 1 done = 0 error = 0}
| org.apache.hadoop.hbase.master.SplitLogManager$TimeoutMonitor.chore
(SplitLogManager.java:745)
2018-03-26 11:11:06,185 | INFO | hadoopc1h3,21300,1522031630949_splitLogManager__ChoreService_1 |
total tasks = 1 unassigned = 0 tasks={/hbase/splitWAL/WALs%2Fhadoopc1h1%2C213
02%2C1520214023667-splitting
%2Fhadoopc1h1%252C21302%252C1520214023667.default.1520584926990=last_update =
1522033841041 last_version = 34255 cur_worker_name = hadoopc1h3,21302,
1520943011826 status = in_progress incarnation = 3 resubmits = 3 batch = installed = 1 done = 0 error = 0}
| org.apache.hadoop.hbase.master.SplitLogManager$TimeoutMonitor.chore
(SplitLogManager.java:745)
2018-03-26 11:11:10,787 | INFO | RpcServer.reader=9,bindAddress=hadoopc1h3,port=21300 | Kerberos
principal name is hbase/hadoop.hadoop.com@HADOOP.COM | org.apache.hadoop.hbase
.ipc.RpcServer$Connection.readPreamble(RpcServer.java:1532)
2018-03-26 11:11:12,185 | INFO | hadoopc1h3,21300,1522031630949_splitLogManager__ChoreService_1 |
total tasks = 1 unassigned = 0 tasks={/hbase/splitWAL/WALs%2Fhadoopc1h1%2C213
02%2C1520214023667-splitting
%2Fhadoopc1h1%252C21302%252C1520214023667.default.1520584926990=last_update =
1522033841041 last_version = 34255 cur_worker_name = hadoopc1h3,21302,
1520943011826 status = in_progress incarnation = 3 resubmits = 3 batch = installed = 1 done = 0 error = 0}
| org.apache.hadoop.hbase.master.SplitLogManager$TimeoutMonitor.chore
(SplitLogManager.java:745)
2018-03-26 11:11:18,185 | INFO | hadoopc1h3,21300,1522031630949_splitLogManager__ChoreService_1 |
total tasks = 1 unassigned = 0 tasks={/hbase/splitWAL/WALs%2Fhadoopc1h1%2C213
02%2C1520214023667-splitting
%2Fhadoopc1h1%252C21302%252C1520214023667.default.1520584926990=last_update =
1522033841041 last_version = 34255 cur_worker_name = hadoopc1h3,21302,
1520943011826 status = in_progress incarnation = 3 resubmits = 3 batch = installed = 1 done = 0 error = 0}
| org.apache.hadoop.hbase.master.SplitLogManager$TimeoutMonitor.chore
(SplitLogManager.java:745)
```

The WAL splitting of RegionServer fails when the node is powered on and off.

## Solution

**Step 1** Stop HBase.

**Step 2** Run the **hdfs fsck** command to check the health status of the **/hbase/WALs** file.

```
hdfs fsck /hbase/WALs
```

If the following command output is displayed, all files are normal. If any file is abnormal, rectify the fault, and then perform the subsequent operations.

```
The filesystem under path '/hbase/WALs' is HEALTHY
```

**Step 3** Back up the **/hbase/WALs** file.

```
hdfs dfs -mv /hbase/WALs /hbase/WALs_old
```

**Step 4** Run the following command to create the **/hbase/WALs** directory.

```
hdfs dfs -mkdir /hbase/WALs
```

Make sure that the permission on the directory is **hbase:hadoop**.

**Step 5** Start HBase.

```
----End
```

## 11.17 Failed to Import HBase Data Due to Oversized File Blocks

### Symptom

Error Message "NotServingRegionException" is displayed when data is imported to HBase.

### Cause Analysis

When a block is greater than 2 GB, a read exception occurs during the seek operation of the HDFS. A full GC occurs when data is frequently written to the RegionServer. As a result, the heartbeat between the HMaster and RegionServer becomes abnormal, and the HMaster marks the RegionServer as dead, and the RegionServer is forcibly restarted. After the restart, the servercrash mechanism is triggered to roll back WALs. Currently, the **splitwal** file has reached 2.1 GB and has only one block. As a result, the HDFS seek operation becomes abnormal and the WAL file splitting fails. However, the RegionServer detects that the WAL needs to be split and triggers the splitwal mechanism, causing a loop between WAL splitting and the splitting failure. In this case, the regions on the RegionServer node cannot be brought online, and an exception is thrown indicating that the region is not online when a region on the RegionServer is queried.

### Procedure

**Step 1** Go to the HBase service page.

- For versions earlier than MRS 3.x. Click the cluster name on the MRS console and choose **Components > HBase**.

 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 3.x or later: Log in to FusionInsight Manager and choose **Cluster**. Click the name of the desired cluster, and choose **Services > HBase**.

**Step 2** On the right of **HMaster Web UI**, click **HMaster (Active)** to go to the HBase Web UI page.

**Step 3** On the **Procedures** page, view the node where the problem occurs.

**Step 4** Log in to the faulty node as user **root** and run the **hdfs dfs -ls** command to view all block information.

**Step 5** Run the **hdfs dfs -mkdir** command to create a directory for storing faulty blocks.

**Step 6** Run the **hdfs dfs -mv** command to move the faulty block to the new directory.

----End

## Summary and Suggestions

The following is provided for your reference:

- If data blocks are corrupted, run the **hdfs fsck /tmp -files -blocks -racks** command to check the health information about data blocks.
- If you perform data operations when a region is being split, **NotServingRegionException** is thrown.

## 11.18 Failed to Load Data to the Index Table After an HBase Table Is Created Using Phoenix

### Symptom

A user fails to run commands to load data to the index table after creating an HBase table using Phoenix. The following error information is displayed:

- MRS 2.x or earlier: Mutable secondary indexes must have the `hbase.regionserver.wal.codec` property set to `org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec` in the `hbase-sites.xml` of every region server. `tableName=MY_INDEX (state=42Y88,code=1029)`

```
Error: ERROR 1029 (42Y88): Mutable secondary indexes must have the hbase.regionserver.wal.codec property set to org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec in the hbase-sites.xml of every region server. tableName=MY_INDEX (state=42Y88,code=1029)
java.sql.SQLException: ERROR 1029 (42Y88): Mutable secondary indexes must have the hbase.regionserver.wal.codec property set to org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec in the hbase-sites.xml of every region server. tableName=MY_INDEX
    at org.apache.phoenix.exception.SQLExceptionCodesFactory$1.newException(SQLException.java:498)
    at org.apache.phoenix.exception.SQLExceptionInfo.buildException(SQLExceptionInfo.java:150)
    at org.apache.phoenix.schema.MetadataClient.createIndex(MetadataClient.java:1534)
    at org.apache.phoenix.compile.CreateIndexCompilers$1.execute(CreateIndexCompiler.java:85)
    at org.apache.phoenix.jdbc.PhoenixStatement$2.call(PhoenixStatement.java:410)
    at org.apache.phoenix.jdbc.PhoenixStatement$2.call(PhoenixStatement.java:393)
    at org.apache.phoenix.call.CallRunner.run(CallRunner.java:53)
    at org.apache.phoenix.jdbc.PhoenixStatement.executeMutation(PhoenixStatement.java:392)
    at org.apache.phoenix.jdbc.PhoenixStatement.executeMutation(PhoenixStatement.java:380)
    at org.apache.phoenix.jdbc.PhoenixStatement.execute(PhoenixStatement.java:1829)
    at sqlline.Commands.execute(Commands.java:822)
    at sqlline.Commands.sql(Commands.java:732)
    at sqlline.SqlLine.dispatch(SqlLine.java:813)
    at sqlline.SqlLine.begin(SqlLine.java:588)
    at sqlline.SqlLine.start(SqlLine.java:308)
    at sqlline.SqlLine.main(SqlLine.java:291)
0: jdbc:phoenix:node-master1@192.168.1.101:2188>
```

- MRS 3. x or later: Exception in thread "main" java.io.IOException: Retry attempted 10 times without completing, bailing out

```
2022-04-17 20:24:37,157 INFO [main] tool.LoadIncrementalHFiles: Split occurred while grouping HFiles, retry attempt 10 with 1 files remaining to group or split
2022-04-17 20:24:37,170 ERROR [main] tool.LoadIncrementalHFiles: .....
Bulk load aborted with some files not yet loaded:
.....
hdfs://hacluster/tmp/3cdc8475-3867-4d9f-a774-87bc6759ee77/ANALYSIS_USER_IDENTIFICATION/F/36b9e9618d784ccf9d982ced4ebb4b76

Exception in thread "main" java.io.IOException: Retry attempted 10 times without completing, bailing out
    at org.apache.hadoop.hbase.tool.LoadIncrementalHFiles.performBulkLoad(LoadIncrementalHFiles.java:460)
    at org.apache.hadoop.hbase.tool.LoadIncrementalHFiles.doBulkLoad(LoadIncrementalHFiles.java:379)
    at org.apache.hadoop.hbase.tool.LoadIncrementalHFiles.doBulkLoad(LoadIncrementalHFiles.java:293)
    at org.apache.phoenix.mapreduce.AbstractBulkLoadTool.completeBulkLoad(AbstractBulkLoadTool.java:389)
    at org.apache.phoenix.mapreduce.AbstractBulkLoadTool.submitJob(AbstractBulkLoadTool.java:343)
    at org.apache.phoenix.mapreduce.AbstractBulkLoadTool.loadData(AbstractBulkLoadTool.java:279)
    at org.apache.phoenix.mapreduce.AbstractBulkLoadTool.run(AbstractBulkLoadTool.java:180)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:90)
    at org.apache.phoenix.mapreduce.JsonBulkLoadTool.main(JsonBulkLoadTool.java:51)
[root@node-master1typi ~]#
```

## Procedure

**Step 1** For MRS 2.x or earlier, perform the following operations:

1. Log in to MRS Manager as user **admin**, choose **Services**, and click **HBase**. On the **Service Configuration** tab, select **All** from the **Type** drop-down list, choose **HMaster > Customization**, and add a configuration item for parameter **hbase.hmaster.config.expandor** with name **hbase.regionserver.wal.codec** and value **org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec**.
2. Choose **RegionServer > Customization**, add a configuration item for parameter **hbase.regionserver.config.expandor** with name **hbase.regionserver.wal.codec** and value **org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec**, and click **Save Configuration**. Then enter the password of the current user and click **OK**.
3. On the **Service Status** page, click **More** and select **Restart Service**. Enter the password of the current user and click **OK** to restart the HBase service.

**Step 2** For MRS 3.x or later, perform the following operations:

1. Log in to FusionInsight Manager as user **admin** and choose **Cluster > Services > HBase**. On the HBase page, choose **Configurations > All Configurations > RegionServer > Customization**. In the right pane, add a configuration item for parameter **hbase.regionserver.config.expandor** with name **hbase.regionserver.wal.codec** and value **org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec**.
2. Choose **HMaster > Customization**, and add a configuration item for parameter **hbase.hmaster.config.expandor** with name **hbase.regionserver.wal.codec** and value **org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec**.
3. Click **Save**. In the dialog box that is displayed, click **OK** to save the configuration.
4. On the **Dashboard** page, click **More** and select **Restart Service**. Enter the password of the current user and click **OK** to restart the HBase service.

----End







## 11.21 HBase Failed to Start Due to Insufficient RegionServer Memory

### Issue

The HBase service fails to start because the remaining RegionServer memory is insufficient.

### Cause Analysis

The troubleshooting process is as follows:

1. Log in to the master node, go to the **/var/log/Bigdata** directory, and search for the HBase log. The log contains error message "connect regionserver timeout".
2. Log in to the RegionServer node in **1** that cannot be connected to HMaster and go to the **/var/log/Bigdata** directory to search for the HBase log. The RegionServer reports error message "error='Cannot allocate memory'(errno=12)".
3. According to the error message in **2**, the startup failure is caused by insufficient RegionServer memory.

### Procedure

**Step 1** Log in to the RegionServer node where the error is reported and run the following command to check the remaining memory of the node:

```
free -g
```

**Step 2** Run the **top** command to check the memory usage of the node.

**Step 3** Stop the memory-consuming processes (not the processes of the MRS components) as prompted and restart the HBase service.

#### NOTE

Besides MRS components, jobs on Yarn are allocated to core nodes in the cluster, thereby occupying node memory. If the startup failure is caused by memory-consuming Yarn jobs, you are advised to expand the capacity of core nodes.

----End

## 11.22 Failed to Start HRegionServer on the Node Newly Added to the Cluster

### Symptom

- After the cluster is scaled out, HRegionserver on the new node fails to be started and remains abnormal.
- Log in to the node where the faulty RegionServer resides and run the **jps** command. The command output shows that the RegionServer process is not

started. Then manually restart the faulty RegionServer instance on Manager, but the restart fails. Check the `/var/log/Bigdata/hbase/rs/hbase-omm-regionserver-node-ana-coreqRvt.log` file of the RegionServer node, and error message "ClassNotFound: org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec" is displayed.

- The following custom configuration of RegionServer is available on Manager.

| Parameter                          | Value                                                                                                                                                                                     | Description | Parameter File |                             |                            |  |  |
|------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|----------------|-----------------------------|----------------------------|--|--|
| HBase->RegionServer                |                                                                                                                                                                                           |             |                |                             |                            |  |  |
| hbase.regionserver.config.expandor | <table border="1"> <thead> <tr> <th>Name</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>hbase.regionserver.wal.code</td> <td>org.apache.hadoop.hbase.re</td> </tr> </tbody> </table> | Name        | Value          | hbase.regionserver.wal.code | org.apache.hadoop.hbase.re |  |  |
| Name                               | Value                                                                                                                                                                                     |             |                |                             |                            |  |  |
| hbase.regionserver.wal.code        | org.apache.hadoop.hbase.re                                                                                                                                                                |             |                |                             |                            |  |  |

## Cause Analysis

The indexing function of Phoenix is configured, but the new node does not have the JAR package of Phoenix. As a result, the related class cannot be found and the startup fails.

## Procedure

**Step 1** Log in to a normal RegionServer node and run the following command:

```
grep -Rn
'org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec' /opt/
Bigdata/MRS_Current/1_16_RegionServer/
```

Two Phoenix packages (**phoenix-4.14.1-server.jar** and **phoenix-core-4.14.1.jar**) in the `/opt/Bigdata/MRS_Current/1_16_RegionServer/install/hbase/lib` directory contain this class.

**Step 2** Run the `scp` command to copy the preceding two packages to the same directory on the faulty RegionServer node and restart the faulty node.

----End

## 11.23 Region in the RIT State for a Long Time Due to HBase File Loss

### Issue

After data is migrated to a new cluster, the region status of HBase is **RIT** or **FAILED\_OPEN**.

### Cause Analysis

1. Log in to the core node of the cluster and check the RegionServer log. The following errors can be found:
  - Failed to open region **3b3ae24c65fc5094bc2acfebaa7a56de**, and the region status is **FAILED\_OPEN**.
  - A file in region **b7b3faab86527b88a92f2a248a54d3dc** cannot be found and error **FileNotFoundException** is displayed.
2. Run the `hbase hbck` command on the HBase client to check the region. The following error message is displayed:

```
ERROR: Found lingering reference file hdfs://hacluster/hbase/news_user_actions/  
3b3ae24c65fc5094bc2acfebaa7a56de/meta/  
0f47cda55fa44cf9aa2599079894aed6.b7b3faab86527b88a92f2a248a54d3dc
```

According to the reference file, region **3b3ae24c65fc5094bc2acfebaa7a56de** is a child region of region **b7b3faab86527b88a92f2a248a54d3dc** and has referenced the file of its parent region. However, the file of the parent region is lost during the migration, and the child region cannot find the reference file.

This exception cannot be resolved by retries. Therefore, the region is in the RIT state for a long period.

## Procedure

- Step 1** On the HDFS client, run the following command to delete the residual file queried in [2](#):

```
hdfs dfs -rmr HDFS path of the residual file
```

Example:

```
hdfs dfs -rmr /hbase/news_user_actions/3b3ae24c65fc5094bc2acfebaa7a56de/  
meta/0f47cda55fa44cf9aa2599079894aed6.b7b3faab86527b88a92f2a248a54d3dc
```

----End

# 12 Using HDFS

---

## 12.1 HDFS NameNode Instances Become Standby After the RPC Port Is Changed

### Issue

After the NameNode RPC port is changed on the page and HDFS is restarted, all NameNodes are in the standby state, causing a cluster exception.

### Symptom

All NameNodes are in the standby state, causing a cluster exception.

### Cause Analysis

After the cluster is installed and started, if the NameNode RPC port is changed, the Zkfc service must be formatted to update node information on ZooKeeper.

### Procedure

**Step 1** Log in to Manager and stop the HDFS service.

 **NOTE**

Do not stop related services when stopping HDFS.

**Step 2** After the services are stopped, log in to the Master node whose RPC port is changed.

 **NOTE**

If the RPC port is changed on both Master nodes, you can log in to either of the Master nodes.

**Step 3** Run the `su - omm` command to switch to user `omm`.

 **NOTE**

For a security cluster, run the `kinit hdfs` command for authentication.

**Step 4** Run the following command to load the environment variable script to the environment:

```
cd ${BIGDATA_HOME}/MRS_X.X.X/1_8_Zkfc/etc
```

```
source ${BIGDATA_HOME}/MRS_X.X.X/install/FusionInsight-Hadoop-3.1.1/  
hadoop/sbin/exportENV_VARS.sh
```

 **NOTE**

In the preceding command, *MRS\_X.X.X* and *1\_8* vary depending on the actual version.

**Step 5** After the loading is complete, run the following command to format the Zkfc:

```
cd ${HADOOP_HOME}/bin  
./hdfs zkfc -formatZK
```

**Step 6** After the formatting is successful, restart HDFS on Manager.

 **NOTE**

If the RPC port of the NameNode is changed, the configuration file must be updated for all clients that have been installed.

----End

## 12.2 An Error Is Reported When the HDFS Client Is Connected Through a Public IP Address

### Symptom

If the host is connected using a public IP address, its HDFS client cannot be used. After the **hdfs** command is executed, message "-bash: hdfs: command not found" is displayed.

### Cause Analysis

The environment variables are not set before the user logs in to the Master node and runs the command.

### Procedure

**Step 1** Log in to any master node as user **root**.

**Step 2** Run the **source *Client installation directory*/bigdata\_env** command to set environment variables.

**Step 3** Run the **hdfs** command to use the HDFS client.

----End

## 12.3 Failed to Use Python to Remotely Connect to the Port of HDFS

### Issue

Failed to use Python to remotely connect to the port of HDFS.

### Symptom

Failed to use Python to remotely connect to port 50070 of HDFS.

### Cause Analysis

The default port of open source HDFS is 50070 for versions earlier than 3.0.0 and is 9870 for version 3.0.0 or later. The port used by the user does not match the HDFS version.

**Step 1** Log in to the active Master node in the cluster.

**Step 2** Run the `su - omm` command to switch to user `omm`.

**Step 3** Run the `/opt/Bigdata/om-0.0.1/sbin/queryVersion.sh` or `sh ${BIGDATA_HOME}/om-server/om/sbin/queryVersion.sh` command to check the HDFS version in the cluster.

Determine the port number of the open-source component based on the version number. For details about how to obtain the port number of an open-source component, see [List of Open Source Component Ports](#).

**Step 4** Run the `netstat -an|grep ${port}` command to check whether the default port number of the component exists.

If it does not exist, the default port number is changed. Change the port to the default port and reconnect to HDFS.

If it exists, contact technical support.

#### NOTE

- `${port}`: indicates the default port number corresponding to the component version.
- If you have changed the default port number, use the new port number to connect to HDFS. You are advised not to change the default port number.

----End

## 12.4 HDFS Capacity Reaches 100%, Causing Unavailable Upper-Layer Services Such as HBase and Spark

### Issue

The HDFS capacity usage of the cluster reaches 100%, and the HDFS service status is read-only. As a result, upper-layer services such as HBase and Spark are unavailable.

### Symptom

The HDFS capacity usage is 100%, the disk capacity usage is only about 85%, and the HDFS service status is read-only. As a result, upper-layer services such as HBase and Spark are unavailable.

### Cause Analysis

Currently, NodeManager and DataNode share data disks. By default, MRS reserves 15% of data disk space for non-HDFS. You can change the percentage of data disk space by setting the HDFS parameter **dfs.datanode.du.reserved.percentage**.

If the HDFS disk usage is 100%, you can set **dfs.datanode.du.reserved.percentage** to a smaller value to restore services and then expand disk capacity.

### Procedure

**Step 1** Log in to any Master node in the cluster.

**Step 2** Run the **source /opt/client/bigdata\_env** command to initialize environment variables.

#### NOTE

If it is a security cluster, run the **kinit -kt <keytab file> <Principal name>** command for authentication.

**Step 3** Run the **hdfs dfs -put ./startDetail.log /tmp** command to check whether HDFS fails to write files.

```
19/05/12 10:07:32 WARN hdfs.DataStreamer: DataStreamer Exception
org.apache.hadoop.ipc.RemoteException(java.io.IOException): File /tmp/startDetail.log_COPYING_ could
only be replicated to 0 nodes instead of minReplication (=1). There are 3 datanode(s) running and no
node(s) are excluded in this operation.
```

**Step 4** Run the **hdfs dfsadmin -report** command to check the used HDFS capacity. The command output shows that the HDFS capacity usage has reached 100%.

```
Configured Capacity: 5389790579100 (4.90 TB)
Present Capacity: 5067618628404 (4.61 TB)
DFS Remaining: 133350196 (127.17 MB)
DFS Used: 5067485278208 (4.61 TB)
DFS Used%: 100.00%
Under replicated blocks: 10
Blocks with corrupt replicas: 0
```



```
Missing blocks: 0  
Missing blocks (with replication factor 1): 0  
Pending deletion blocks: 0
```

**Step 5** When the HDFS capacity usage reaches 100%, change the percentage of data disk space by setting HDFS parameter **dfs.datanode.du.reserved.percentage**.

1. Go to the service configuration page.
  - MRS Manager: Log in to MRS Manager and choose **Services > HDFS > Configuration**.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > HDFS > Configurations**.
2. Click **All Configurations** and search for **dfs.datanode.du.reserved.percentage**.
3. Change the value of this parameter to **10**.

**Step 6** After the modification, increase the number of disks of the Core node.

----End

## 12.5 Error Message "Permission denied" Is Displayed When HDFS and Yarn Are Started

### Issue

An error is reported during HDFS and Yarn startup.

### Symptom

HDFS and Yarn fail to be started. The following error information is displayed: **/dev/null Permission denied**

```
[2018-11-16 08:52:57] Start service 'ServiceName: Yarn'.  
[2018-11-16 08:52:57] Start role 'ROLE[name: ResourceManager]'.  
[2018-11-16 08:52:57] Start role 'ROLE[name: NodeManager]'.  
[2018-11-16 08:52:57] Start role instance 'ResourceManager#192.168.0.23@node-master2-CMCgr'.  
[2018-11-16 08:52:57] Start role instance 'ResourceManager#192.168.0.59@node-master1-bdWZs'.  
[2018-11-16 08:52:57] Start role instance 'NodeManager#192.168.0.37@node-core-gKPas'.  
[2018-11-16 08:52:57] Start role instance 'NodeManager#192.168.0.137@node-core-qFOXf'.  
[2018-11-16 08:52:57] Start role instance 'NodeManager#192.168.0.135@node-core-nDKmI'.  
[2018-11-16 08:52:57] Start the role instance for 'ROLE[name: ResourceManager]' successfully.  
[2018-11-16 08:52:57] Start the role instance for 'ROLE[name: NodeManager]' successfully.  
[2018-11-16 08:52:57] Start the role instance for 'ROLE[name: NodeManager]' successfully.  
[2018-11-16 08:52:57] Start the role instance for 'ROLE[name: NodeManager]' successfully.  
[2018-11-16 08:52:57] Start the role for 'ServiceName: Yarn' successfully.  
Fail to prepare to start role instance 'NodeManager#192.168.0.135@node-core-nDKmI' [ScriptExecutionResult=ScriptExecutionResult [exitCode=1, output=, errMsg=/etc/bashrc: line 84: /dev/null: Permission denied
```

### Cause Analysis

The **/dev/null** permission value of the VM is changed to **775**.

```
70 cd ..
71 ll
72 chmod -R 775 /dev/
73 ll
74 chmod -r 775 dbdata_on/
75 ll
76 chmod -r 770 dbdata_on/
77 ll
78 chmod -r 777 dbdata_on/
79 ll
80 cd ..
81 ll
```

## Procedure

- Step 1** Log in to any master node in the cluster as user **root**.
- Step 2** After successful login, run the **chmod 666 /dev/null** command to change the permission value of **/dev/null** to **666**.
- Step 3** Run the **ls -al /dev/null** command to check whether the new permission value of **/dev/null** is **666**. If it is not, change the value to **666**.
- Step 4** After the modification is successful, restart HDFS and Yarn.

----End

## 12.6 HDFS Users Can Create or Delete Files in Directories of Other Users

### Symptom

When using MRS, a user has the permission to delete or create files in another user's HDFS directory.

### Cause Analysis

The user has the permission for the **ficommon** group and therefore can perform any operations on the HDFS. You need to remove the user's **ficommon** group permission.

### Procedure

- Step 1** Log in to the master node in the cluster as user **root**.
- Step 2** Run the **id \${Username}** command to check whether the user has the **ficommon** group permission.  
  
If the user has the **ficommon** group permission, go to **Step 3**. If the user does not have the **ficommon** group permission, contact technical support.

#### NOTE

*Username* indicates the name of the user whose HDFS permission is incorrectly set.

- Step 3** Run the **gpasswd -d \${Username} ficommon** command to delete the user's **ficommon** group permission.

**Step 4** Modify parameters on Manager.

MRS Manager (applicable to versions earlier than MRS 3.x):

1. Log in to MRS Manager and choose **Services > HDFS > Service Configuration**.
2. Set **Type** to **All**, enter **dfs.permissions.enabled** in the search box, and change the parameter value to **true**.
3. Click **Save Configuration** and restart the HDFS service.

FusionInsight Manager (applicable to MRS 3.x or later):

1. Log in to FusionInsight Manager. Choose **Cluster > Services > HDFS > Configurations > All Configurations**.
2. Enter **dfs.permissions.enabled** in the search box and change the value to **true**.
3. After the modification is complete, click **Save** and restart the HDFS service.

MRS console :

1. Log in to the MRS console and choose **Components > HDFS > Service Configuration**.
2. Set **Type** to **All**, enter **dfs.permissions.enabled** in the search box, and change the parameter value to **true**.
3. Click **Save Configuration** and restart the HDFS service.

----End

## 12.7 A DataNode of HDFS Is Always in the Decommissioning State

### Issue

A DataNode of HDFS is in the **Decommissioning** state for a long period of time.

### Symptom

A DataNode of HDFS fails to be decommissioned (or the Core node fails to be scaled in), but the DataNode remains in the Decommissioning state.

### Cause Analysis

During the decommissioning of a DataNode (or scale-in of the Core node) in HDFS, the decommissioning or scale-in task fails and the blacklist is not cleared because the Master node is restarted or the NodeAgent process exits unexpectedly. In this case, the DataNode remains in the **Decommissioning** state. The blacklist needs to be cleared manually.

### Procedure

- Step 1** Go to the service instance page.

MRS Manager:

Log in to MRS Manager and choose **Services > HDFS > Instance**.

FusionInsight Manager:

MRS 3.x or later: Log in to FusionInsight Manager and choose **Cluster > Service > HDFS > Instance**.

Log in to the MRS console and choose **Components > HDFS > Instances**.

- Step 2** Check the HDFS service instance status, locate the DataNode that is in the decommissioning state, and copy the IP address of the DataNode.
- Step 3** Log in to the Master1 node and run the `cd ${BIGDATA_HOME}/MRS_*/1_*_NameNode/etc/` command to go to the blacklist directory.
- Step 4** Run the `sed -i "/^IP$/d" excludeHosts` command to clear the faulty DataNode information from the blacklist. Replace the IP address in the command with the IP address of the faulty DataNode queried in [Step 2](#). The IP address cannot contain spaces.
- Step 5** If there are two Master nodes, perform [Step 3](#) and [Step 4](#) on Master2.
- Step 6** Run the following command on the Master1 node to initialize environment variables:
- ```
source Client installation directory/bigdata_env
```
- Step 7** 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, skip this step:
- ```
kinit Service user who has the HDFS operation permission
```
- Step 8** Run the following command on the Master1 node to update the HDFS blacklist:
- ```
hdfs dfsadmin -refreshNodes
```
- Step 9** Run the `hdfs dfsadmin -report` command to check the status of each DataNode. Ensure that the DataNode corresponding to the IP address obtained has been restored to the **Normal** state.

Figure 12-1 DataNode status

```
Name: 192.168.2.230:9866 (node-ana-coreoYfm)
Hostname: node-ana-coreoYfm
Rack: /default/rack0
Decommission Status : Normal
Configured Capacity: 105554829312 (98.31 GB)
DFS Used: 1225715740 (1.14 GB)
Non DFS Used: 3045261284 (2.84 GB)
DFS Remaining: 95361495372 (88.81 GB)
DFS Used%: 1.16%
DFS Remaining%: 90.34%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 10
Last contact: Thu Aug 15 15:53:17 CST 2019
Last Block Report: Thu Aug 15 12:12:46 CST 2019
Num of Blocks: 974
```

**Step 10** Go to the service instance page.

MRS Manager:

Log in to MRS Manager and choose **Services > HDFS > Instances**.

FusionInsight Manager:

MRS 3.x or later: Log in to FusionInsight Manager and choose **Cluster > Service > HDFS > Instance**.

Log in to the MRS console and choose **Components > HDFS > Instances**.

**Step 11** Select the DataNode instance that is in the decommissioning state and choose **More > Restart Instance**.

**Step 12** Wait until the restart is complete and check whether the DataNode is restored.

----End

## Summary and Suggestions

Do not perform high-risk operations, such as restarting nodes, during decommissioning (or scale-in).

## Related Information

None

# 12.8 HDFS NameNode Failed to Start Due to Insufficient Memory

## Symptom

Scenario 1: After the HDFS service is restarted, HDFS is in the **Bad** state, and the NameNode instance status is abnormal and cannot exit the safe mode for a long time.

Scenario 2: The NameNode fails to be started after the startup times out, and the native web UI cannot be opened.

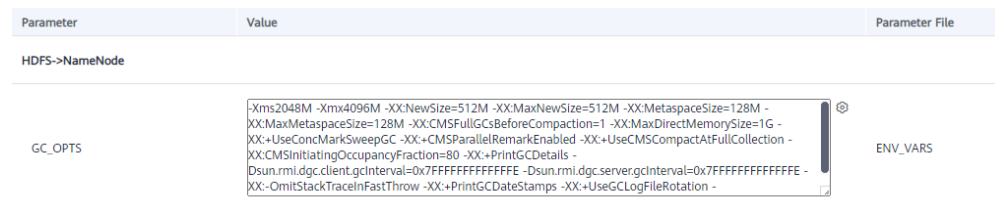
## Cause Analysis

1. In the NameNode run log (**`/var/log/Bigdata/hdfs/nn/hadoop-omm-namenode-XXX.log`**), search for **WARN**. It is found that GC takes 63 seconds.  
2017-01-22 14:52:32,641 | WARN | org.apache.hadoop.util.JvmPauseMonitor\$Monitor@1b39fd82 | Detected pause in JVM or host machine (eg GC): pause of approximately 63750ms  
GC pool 'ParNew' had collection(s): count=1 time=0ms  
GC pool 'ConcurrentMarkSweep' had collection(s): count=1 time=63924ms | JvmPauseMonitor.java:189
2. Analyze the NameNode log **`/var/log/Bigdata/hdfs/nn/hadoop-omm-namendoe-XXX.log`**. It is found that the NameNode is waiting for block reporting and the total number of blocks is too large. In the following example, the total number of blocks is 36.29 million.

2017-01-22 14:52:32,641 | INFO | IPC Server handler 8 on 25000 | STATE\* Safe mode ON.  
The reported blocks 29715437 needs additional 6542184 blocks to reach the threshold 0.9990 of total blocks 36293915.

3. On Manager, check the **GC\_OPTS** parameter of the NameNode:

**Figure 12-2** Checking the GC\_OPTS parameter of the NameNode



4. For details about the mapping between the NameNode memory configuration and data volume, see [Table 12-1](#).

**Table 12-1** Mapping between NameNode memory configuration and data volume

Number of File Objects	Reference Value
10,000,000	<code>-Xms6G -Xmx6G -XX:NewSize=512M -XX:MaxNewSize=512M</code>
20,000,000	<code>-Xms12G -Xmx12G -XX:NewSize=1G -XX:MaxNewSize=1G</code>
50,000,000	<code>-Xms32G -Xmx32G -XX:NewSize=2G -XX:MaxNewSize=3G</code>
100,000,000	<code>-Xms64G -Xmx64G -XX:NewSize=4G -XX:MaxNewSize=6G</code>
200,000,000	<code>-Xms96G -Xmx96G -XX:NewSize=8G -XX:MaxNewSize=9G</code>
300,000,000	<code>-Xms164G -Xmx164G -XX:NewSize=12G -XX:MaxNewSize=12G</code>

## Solution

- Step 1** Modify the NameNode memory parameter based on the specifications. If the number of blocks is 36 million, change the parameter value to **`-Xms32G -Xmx32G -XX:NewSize=2G -XX:MaxNewSize=3G`**.
  - Step 2** Restart a NameNode and check that the NameNode can be started normally.
  - Step 3** Restart the other NameNode and check that the page status is restored.
- End

## 12.9 A Large Number of Blocks Are Lost in HDFS due to the Time Change Using ntpdate

### Symptom

1. A user uses **ntpdate** to change the time for a cluster that is not stopped. After the time is changed, HDFS enters the safe mode and cannot be started.
2. After HDFS exits the safe mode and starts, about 1 TB data is lost during the **hfck** check.

### Cause Analysis

1. A large number of blocks are lost on the native NameNode page.

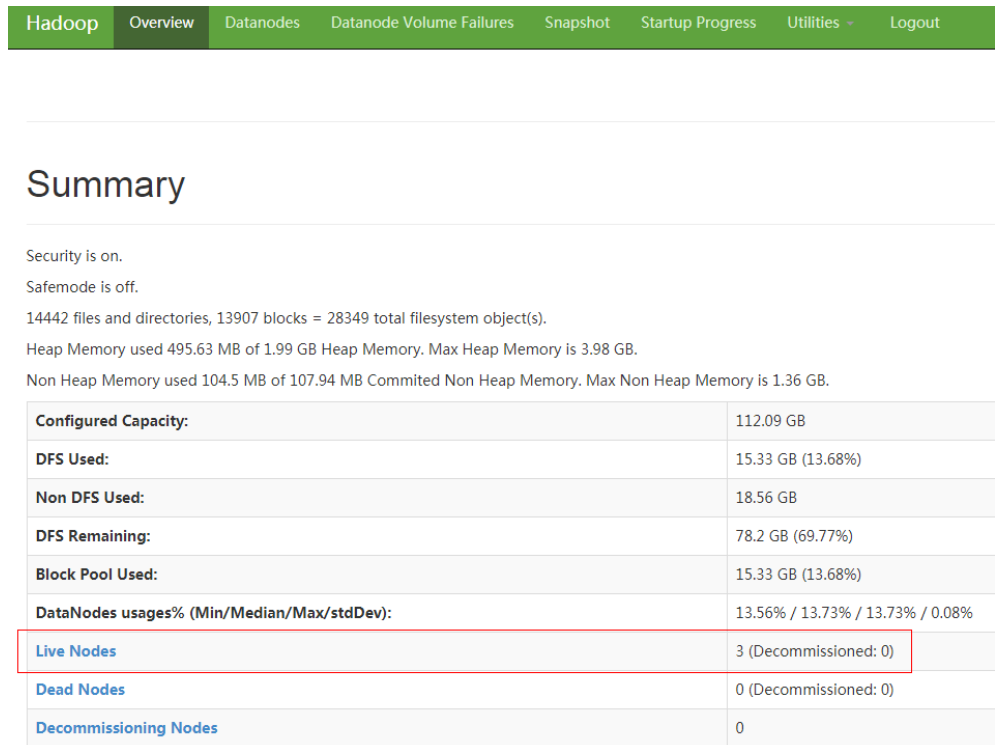
Figure 12-3 Block loss

```
There are 41491 missing blocks. The following files may be corrupted:

blk_1090519588 /user/etlhadoo/struct_data/uds_data/PRS/20180130/DCM_FRS_PDWTMDTL_S_000_input/1/ccw-20180130-pdwtmdtl-023_022_bin_7
blk_1090520189 /user/hive/warehouse/prs_mc.db/dcm_prs_pdwtmdtl_s/pt_dt=2018-01-30/part-m-00004
blk_1082131961 /user/hive/warehouse/cas_mc.db/dcm_cas_nthpatel_h/end_dt=2017-12-31/000004_0
blk_1082132310 /user/hive/warehouse/crl_mc.db/dcm_crl_ecs_tb2045_s/pt_dt=2017-12-31/000005_0
blk_1082132604 /user/hive/warehouse/crl_mc.db/dcm_crl_ecs_tb2045_s/pt_dt=2017-12-31/000040_0
blk_1090521279 /user/hive/warehouse/gcm_mc.db/dcm_gcm_pndllta200211_h/end_dt=2018-01-30/000008_0
blk_1090521294 /user/hive/warehouse/gcm_mc.db/dcm_gcm_pndllta200211_h/end_dt=2018-01-30/000012_0
blk_1090521427 /user/hive/warehouse/pis_mc.db/dcm_pis_lthpedtl_h/end_dt=2018-01-30/000080_0
blk_1090521473 /user/hive/warehouse/pis_mc.db/dcm_pis_lthpedtl_h/end_dt=2018-01-30/000016_0
blk_1082133176 /user/hive/warehouse/cas_mc.db/dcm_cas_hffpbat_s/pt_dt=2017-12-31/part-m-00006
blk_1090522261 /user/etlhadoo/struct_data/uds_data/ECS/20180130/DCM_ECS_TB1170_S_000_input/1/ciw-20180130-hdwbl171-022_032_bin_16
blk_1090522656 /user/etlhadoo/struct_data/uds_data/ECS/20180130/DCM_ECS_TB1170_S_output/1/part-m-00007
blk_1090522747 /user/hive/warehouse/gcm_mc.db/dcm_gcm_rassure_change_detail_s/pt_dt=2018-01-31/000002_0
blk_1082134372 /user/hive/warehouse/bcs_mc.db/dcm_bcs_bthrsism_h/pt_dt=2017-12-31/part-m-00006
blk_1090523585 /user/hive/warehouse/ecs_mc.db/dcm_ecs_tbi170_s/pt_dt=2018-01-30/000002_0
blk_1090523811 /user/hive/warehouse/nae_mc.db/dcm_nae_nfpjnl_s/pt_dt=2018-01-30/part-m-00005
blk_1082135337 /user/hive/warehouse/bcs_mc.db/dcm_bcs_bthrsism_h/pt_dt=2017-12-31/part-m-00022
blk_1090524043 /user/hive/warehouse/nae_mc.db/dcm_nae_nfpjnl_s/pt_dt=2018-01-30/part-m-00016
blk_1082136206 /user/hive/warehouse/bcs_mc.db/dcm_bcs_bthrsism_h/pt_dt=2017-12-31/part-m-00038
blk_1090525355 /user/hive/warehouse/bdsp_bcas_act.db/bcs_jzcs_detail/pt_dt=2017-11-30/000006_0
blk_1090526191 /user/hive/warehouse/bdsp_bcas_act.db/bcs_jzcs_detail/pt_dt=2017-11-30/000008_0
blk_1090526995 /user/hive/warehouse/bdsp_bcas_act.db/bcs_jzcs_detail/pt_dt=2017-11-30/000014_0
blk_1082140552 /user/hive/warehouse/co8_mc.db/m01_co8_corp_cust_mgr/pt_dt=2017-12-31/000001_0
blk_1090529399 /user/hive/warehouse/bdsp_bcas_act.db/bcs_jzcs_middle_t/pt_dt=2017-11-30/000017_0
blk_1090529420 /user/hive/warehouse/bdsp_bcas_act.db/bcs_jzcs_middle_t/pt_dt=2017-11-30/000014_0
blk_1082141596 /user/hive/warehouse/asa_mc.db/t80_asa_bcas_agt_stat/pt_dt=2017-12-31/000032_0
blk_1082141631 /user/hive/warehouse/asa_mc.db/t80_asa_bcas_agt_stat/pt_dt=2017-12-31/000003_0
blk_1082142345 /user/hive/warehouse/zum_mc.db/e00_prod_level_overview_h/pt_dt=2017-12-31/000000_0_copy_1514441562192
blk_1090531076
/user/etlhadoo/struct_data/uds_data/GCM/20180131/DCM_GCM_DEDUM_STOP_PARA_S_000_input/1/CMA_DEDUM_STOP_PARA011800000-011-20180131.REM_11_UTF
blk_1090531330 /user/hive/warehouse/gcc_mc.db/dcm_gcc_rcorp_motor_info_s/pt_dt=2018-01-31/000011_0
blk_1090531342 /user/hive/warehouse/gcc_mc.db/dcm_gcc_rcorp_motor_info_s/pt_dt=2018-01-31/000002_0
blk_1090531494
/user/etlhadoo/struct_data/uds_data/GCM/20180131/DCM_GCM_ZMORTGAGE_PROJECT_STAT_S_000_input/1/CMA_ZMORTGAGE_PROJECT_STAT0050100000-
```

2. DataNode information on the native page shows that the number of displayed DataNodes is 10 less than that of actual DataNodes.

**Figure 12-4** Checking the number of DataNodes



3. Check the DataNode run log file `/var/log/Bigdata/hdfs/dn/hadoop-omm-datanode-hostname.log`. The following error information is displayed:  
Major error information: Clock skew too great

**Figure 12-5** DataNode run log error

```

at org.apache.hadoop.ipc.Client.call(Client.java:1486)
at org.apache.hadoop.ipc.Client.call(Client.java:1447)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:229)
at com.sun.proxy.$Proxy14.versionRequest(Unknown Source)
at org.apache.hadoop.hdfs.protocolPB.DatanodeProtocolClientSideTranslatorPB.versionRequest(DatanodeProtocolClientSideTranslatorPB.java:273)
at org.apache.hadoop.hdfs.server.datanode.BFSerivceActor.retrieveNamespaceInfo(BFSerivceActor.java:187)
at org.apache.hadoop.hdfs.server.datanode.BFSerivceActor.connectToNNAndHandshake(BFSerivceActor.java:237)
at org.apache.hadoop.hdfs.server.datanode.BFSerivceActor.run(BFSerivceActor.java:689)
at java.lang.Thread.run(Thread.java:745)
Caused by: GSSException: No valid credentials provided (Mechanism level: Clock skew too great (37))
at sun.security.jgss.krb5.Krb5Context.initSecContext(Krb5Context.java:770)
at sun.security.jgss.GSSContextImpl.initSecContext(GSSContextImpl.java:248)
at sun.security.jgss.GSSContextImpl.initSecContext(GSSContextImpl.java:179)
at com.sun.security.sasl.gsskerb.GssKrb5Client.evaluateChallenge(GssKrb5Client.java:192)
... 20 more
Caused by: KrbException: Clock skew too great (37)
at sun.security.krb5.KrbKdcRep.check(KrbKdcRep.java:88)
at sun.security.krb5.KrbTgsRep.<init>(KrbTgsRep.java:87)
at sun.security.krb5.KrbTgsReq.getReply(KrbTgsReq.java:259)
at sun.security.krb5.KrbTgsReq.sendAndGetCreds(KrbTgsReq.java:270)
at sun.security.krb5.internal.CredentialsUtil.serviceCreds(CredentialsUtil.java:302)
at sun.security.krb5.internal.CredentialsUtil.acquireServiceCreds(CredentialsUtil.java:120)
at sun.security.krb5.Credentials.acquireServiceCreds(Credentials.java:458)
at sun.security.jgss.krb5.Krb5Context.initSecContext(Krb5Context.java:693)

```

## Solution

**Step 1** Change the time of the 10 DataNodes that cannot be viewed on the native page.

**Step 2** On Manager, restart the DataNode instances.

----End



# 12.10 CPU Usage of DataNodes Is Close to 100% Occasionally, Causing Node Loss

## Symptom

There is a possibility that the CPU usage of DataNodes is close to 100%. As a result, nodes may be lost (the SSH connection is slow or fails).

Figure 12-6 DataNode CPU usage close to 100%

PID	USER	PP	NI	VIRT	RES	SHR	S	SCPU	MEM	TIME+	COMMAND
60636	nnm	20	0	2415m	179	18m	R	279	0.3	1522.08	/opt/csw/bin/psccs -type datanode -mountfile /var/log/hadoop/hdfs/nn/psccs -errfile /var/log/hadoop/hdfs/nn/psccs -pidfile /var/log/hadoop/hdfs/nn/psccs -pidfile
02428	oosadm	20	0	18116	3764	1828	R	155	0.0	1:17.63	/opt/osp/manager/ctsp/python/bin/python /opt/osp/manager/agent-1.3.10.200/tools/psyscript/psyscript.py --cmd status --te
02410	oosadm	20	0	55016	8048	2836	R	155	0.0	1:59.80	/opt/osp/manager/ctsp/python/bin/python /opt/osp/manager/agent-1.3.10.200/tools/psyscript/watchdog.py --cmd status --te
02412	oosadm	20	0	24752	5912	2340	R	155	0.0	1:50.22	/opt/osp/manager/ctsp/python/bin/python /opt/osp/manager/agent-1.3.10.200/tools/psyscript/psyscript.py --cmd psccsinfo --
02484	nnm	20	0	12800	1476	1124	R	155	0.0	0:10.73	/bin/bash -c /jdk1.7.0_80/bin/java -server -Xms1024m -Djava.io.tmpdir=/export/data1/yarn/nn/localdi
02341	oosadm	20	0	57760	8688	3000	R	139	0.0	3:29.41	/opt/osp/manager/ctsp/python/bin/python /opt/osp/manager/agent/tools/psyscript/psyscript.py --cmd status --te
02531	nnm	20	0	11176	640	468	R	106	0.0	0:04.13	/bin/bash -c echo \$NAME \$NN \$NN

## Cause Analysis

- 1. A lot of write failure logs exist on DataNodes.

Figure 12-7 DataNode write failure log

```
2015-08-31 11:29:34,184 | ERROR | DataXceiver for client DFSClient_NONMAPREDUCE_1675952887_23 at /192.168.8.40:44514 [Receiving block BP-125271511-192.168.8.29-1440656260530:blk_1074766997_1034914] | TSP21:25009:DataXceiver error processing WRITE_BLOCK operation src: /192.168.8.40:44514 dst: /192.168.8.64:25009 | DataXceiver.java:258
java.io.IOException: Premature EOF from inputStream
    at org.apache.hadoop.io.IOUtils.readFully(IOUtils.java:194)
    at org.apache.hadoop.hdfs.protocol.datatransfer.PacketReceiver.doReadFully(PacketReceiver.java:213)
    at org.apache.hadoop.hdfs.protocol.datatransfer.PacketReceiver.doRead(PacketReceiver.java:134)
    at org.apache.hadoop.hdfs.protocol.datatransfer.PacketReceiver.receiveNextPacket(PacketReceiver.java:109)
    at org.apache.hadoop.hdfs.server.datanode.BlockReceiver.receivePacket(BlockReceiver.java:446)
    at org.apache.hadoop.hdfs.server.datanode.BlockReceiver.receiveBlock(BlockReceiver.java:707)
    at org.apache.hadoop.hdfs.server.datanode.DataXceiver.writeBlock(DataXceiver.java:748)
    at org.apache.hadoop.hdfs.protocol.datatransfer.Receiver.opWriteBlock(Receiver.java:124)
    at org.apache.hadoop.hdfs.protocol.datatransfer.Receiver.processOp(Receiver.java:71)
    at org.apache.hadoop.hdfs.server.datanode.DataXceiver.run(DataXceiver.java:240)
    at java.lang.Thread.run(Thread.java:745)
2015-08-31 11:29:35,147 | INFO | DataXceiver for client DFSClient_NONMAPREDUCE_402997805_1 at /192.168.8.30:59449 [Sending block BP-125271511-192.168.8.29-1440656260530:blk_1074181856_446655] | src: /192.168.8.64:25009, dest: /192.168.8.30:59449, bytes: 16826, op: HDFS_READ, cliID: DFSClient_NONMAPREDUCE_402997805_1, offset: 0, srvID: 9d1d30a5-046d-438b-83c9-2c6c54c6bd12, blockid: BP-125271511-192.168.8.29-1440656260530:blk_1074181856_446655, duration: 78832 | BlockSender.java:738
2015-08-31 11:29:35,269 | INFO | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 7480ms
No GCs detected | JvmPauseMonitor.java:172
2015-08-31 11:29:36,965 | INFO | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 1215ms
No GCs detected | JvmPauseMonitor.java:172
2015-08-31 11:29:43,067 | INFO | DataXceiver for client DFSClient_NONMAPREDUCE_1675952887_23 at /192.168.8.33:35530 [Receiving block BP-125271511-192.168.8.29-1440656260530:blk_1074767006_1034923] | Exception for BP-125271511-192.168.8.29-1440656260530:blk_1074767006_1034923 | BlockReceiver.java:742
java.io.IOException: Premature EOF from inputStream
```

- 2. A large number of files are written in a short time, causing insufficient DataNode memory.

Figure 12-8 Insufficient DataNode memory

```
Line 153101: 2015-08-31 11:24:29,313 | INFO | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 1199ms
Line 153102: 2015-08-31 11:24:42,689 | WARN | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 11273ms
Line 153103: 2015-08-31 11:24:45,810 | INFO | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 1005ms
Line 153104: 2015-08-31 11:24:49,801 | INFO | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 1067ms
Line 153105: 2015-08-31 11:25:10,167 | WARN | org.apache.hadoop.util.JvmPauseMonitor$Monitor@551bd2a0 | Detected pause in JVM or host machine (eg GC): pause of approximately 12323ms
```

## Solution

**Step 1** Check DataNode memory configuration and whether the remaining server memory is sufficient.

**Step 2** Increase DataNode memory and restart the DataNode.

----End

## 12.11 Manually Performing Checkpoints When a NameNode Is Faulty for a Long Time

### Symptom

If the standby NameNode is faulty for a long time, a large amount of editlog will be accumulated. In this case, if the HDFS or active NameNode is restarted, the active NameNode reads a large amount of unmerged editlog. As a result, the HDFS or active NameNode takes a long time to restart and even fails to restart.

### Cause Analysis

The standby NameNode periodically combines editlog files and generates the fsimage file. This process is called checkpoint. After the fsimage file is generated, the standby NameNode transfers it to the active NameNode.

#### NOTE

As the standby NameNode periodically combines editlog files, it cannot combine them when it becomes abnormal. As a result, the active NameNode needs to load many editlog files during its next startup, which occupies much memory and takes a long time.

The period of metadata combination is determined by the following parameters. If the NameNode runs for 30 minutes or one million counts of operations are performed on HDFS, the checkpoint is implemented.

- **dfs.namenode.checkpoint.period**: specifies the checkpoint period. The default value is **1800s**.
- **dfs.namenode.checkpoint.txns**: specifies the times of operations for triggering the checkpoint execution. The default value is **1000000**.

### Solution

Before restarting the HDFS or active NameNode, perform checkpoint manually to merge metadata of the active NameNode.

**Step 1** Stop workloads.

**Step 2** Obtain the hostname of the active NameNode.

**Step 3** Run the following commands on the client:

```
source /opt/client/bigdata_env
```

```
kinit Component user
```

Note: Replace **/opt/client** with the actual installation path of the client.

**Step 4** Run the following command to enable the safe mode for the active NameNode (replace **linux22** with the hostname of the active NameNode):

```
hdfs dfsadmin -fs linux22:25000 -safemode enter
```

```
linux16:/opt/ki_client # hdfs dfsadmin -fs linux22:25000 -safemode enter
17/04/26 18:38:30 WARN fs.FileSystem: "linux22:25000" is a deprecated filesystem name. Use "hdfs://linux22:25000/" instead.
17/04/26 18:38:32 INFO hdfs.PeerCache: SocketCache disabled.
Safe mode is ON
```

**Step 5** Run the following command to merge editlog on the active NameNode:

```
hdfs dfsadmin -fs linux22:25000 -saveNamespace
```

```
linux16:/opt/Fl_client # hdfs dfsadmin -fs linux22:25000 -saveNamespace
17/04/26 18:38:54 WARN fs.FileSystem: "linux22:25000" is a deprecated filesystem name. Use "hdfs://linux22:25000/" instead.
17/04/26 18:38:56 INFO hdfs.PeerCache: SocketCache disabled.
Save namespace successful
```

**Step 6** Run the following command to make the active NameNode exit the safe mode:

```
hdfs dfsadmin -fs linux22:25000 -safemode leave
```

```
linux16:/opt/Fl_client # hdfs dfsadmin -fs linux22:25000 -safemode leave
17/04/26 18:39:07 WARN fs.FileSystem: "linux22:25000" is a deprecated filesystem name. Use "hdfs://linux22:25000/" instead.
17/04/26 18:39:09 INFO hdfs.PeerCache: SocketCache disabled.
Safe mode is OFF
```

**Step 7** Check whether the combination is complete.

```
cd /srv/BigData/namenode/current
```

Check whether the time of the first generated fsimage is the current time. If yes, the combination is complete.

```
rw----- 1 omm wheel 29447 Apr 26 16:42 edits_000000000002002021_000000000002003017
-rw----- 1 omm wheel 1048576 Apr 26 18:43 edits_inprogress_000000000002083018
-rw----- 1 omm wheel 736657 Apr 26 15:46 fsimage_0000000000002071390
-rw----- 1 omm wheel 62 Apr 26 15:46 fsimage_0000000000002071390.md5
-rw----- 1 omm wheel 736657 Apr 26 16:46 fsimage_0000000000002075405
-rw----- 1 omm wheel 62 Apr 26 16:46 fsimage_0000000000002075405.md5
-rw----- 1 omm wheel 736410 Apr 26 17:46 fsimage_0000000000002079398
-rw----- 1 omm wheel 62 Apr 26 17:46 fsimage_0000000000002079398.md5
-rw----- 1 omm wheel 8 Apr 26 18:42 seen_txid
linux-20:/srv/BigData/namenode/current #
linux-20:/srv/BigData/namenode/current #
```

----End

## 12.12 Error "Failed to place enough replicas" Is Reported When HDFS Reads or Writes Files

### Symptom

When a user performs a write operation on HDFS, the message "Failed to place enough replicas:expected..." is displayed.

### Cause Analysis

- The data receiver of the DataNode is unavailable.

The DataNode log is as follows:

```
2016-03-17 18:51:44,721 | WARN |
org.apache.hadoop.hdfs.server.datanode.DataXceiverServer@5386659f |
hadoopc1h2:25009:DataXceiverServer: | DataXceiverServer.java:158
java.io.IOException: Xceiver count 4097 exceeds the limit of concurrent xcievers: 4096
at org.apache.hadoop.hdfs.server.datanode.DataXceiverServer.run(DataXceiverServer.java:140)
at java.lang.Thread.run(Thread.java:745)
```

- The disk space configured for the DataNode is insufficient.
- DataNode heartbeats are delayed.

### Solution

- If the DataNode data receiver is unavailable, add the value of the HDFS parameter **dfs.datanode.max.transfer.threads** on Manager.

- If disk space or CPU resources are insufficient, add DataNodes or ensure that disk space and CPU resources are available.
- If the network is faulty, ensure that the network is available.

## 12.13 Maximum Number of File Handles Is Set to a Too Small Value, Causing File Reading and Writing Exceptions

### Symptom

The maximum number of file handles is set to a too small value, causing insufficient file handles. Writing files to HDFS is slow or file writing fails.

### Cause Analysis

1. The DataNode log `/var/log/Bigdata/hdfs/dn/hadoop-omm-datanode-XXX.log` contains exception information "java.io.IOException: Too many open files."  

```
2016-05-19 17:18:59,126 | WARN |  
org.apache.hadoop.hdfs.server.datanode.DataXceiverServer@142ff9fa |  
YSDN12:25009:DataXceiverServer: |  
org.apache.hadoop.hdfs.server.datanode.DataXceiverServer.run(DataXceiverServer.java:160)  
java.io.IOException: Too many open files  
at sun.nio.ch.ServerSocketChannellImpl.accept0(Native Method)  
at sun.nio.ch.ServerSocketChannellImpl.accept(ServerSocketChannellImpl.java:241)  
at sun.nio.ch.ServerSocketAdaptor.accept(ServerSocketAdaptor.java:100)  
at org.apache.hadoop.hdfs.net.TcpPeerServer.accept(TcpPeerServer.java:134)  
at org.apache.hadoop.hdfs.server.datanode.DataXceiverServer.run(DataXceiverServer.java:137)  
at java.lang.Thread.run(Thread.java:745)
```
2. The error indicates insufficient file handles. File handles cannot be opened and data is written to other DataNodes. As a result, writing files is slow or fails.

### Solution

- Step 1** Run the `ulimit -a` command to check the maximum number of file handles set for the involved node. If the value is small, change it to **640000**.

**Figure 12-9** Check the number of file handles.

```
[omm@189-39-150-167 ~]$ ulimit -a
core file size          (blocks, -c) 0
data seg size          (kbytes, -d) unlimited
scheduling priority    (-e) 0
file size              (blocks, -f) unlimited
pending signals        (-i) 256551
max locked memory      (kbytes, -l) 64
max memory size        (kbytes, -m) unlimited
open files             (-n) 640000
pipe size              (512 bytes, -p) 8
POSIX message queues   (bytes, -q) 819200
real-time priority     (-r) 0
stack size             (kbytes, -s) 10240
cpu time               (seconds, -t) unlimited
max user processes     (-u) 60000
virtual memory         (kbytes, -v) unlimited
file locks             (-x) unlimited
```

- Step 2** Run the `vi /etc/security/limits.d/90-nofile.conf` command to edit the file and change the number of file handles. If the file does not exist, create one and modify the file as follows:

**Figure 12-10** Changing the number of file handles

```
*          hard    nofile    640000
*          soft    nofile    640000
~
```

- Step 3** Open another terminal. Run the `ulimit -a` command to check whether the modification is successful. If the modification fails, perform the preceding operations again.
- Step 4** Restart the DataNode instance on Manager.

----End

## 12.14 HDFS Client File Fails to Be Closed After Data Writing

### Symptom

An HDFS client file fails to be closed after data is written to the file. A message is displayed indicating that the data block does not have enough replicas.

Client log:

```
2015-05-27 19:00:52.811 [pool-2-thread-3] ERROR: /tsp/nedata/collect/UGW/ugwufdr/
20150527/10/6_20150527105000_20150527105500_SR5S14_1432723806338_128_11.pkg.tmp143272380633
8 close hdfs sequence file fail (SequenceFileInfoChannel.java:444)
java.io.IOException: Unable to close file because the last block does not have enough number of replicas.
at org.apache.hadoop.hdfs.DFSOutputStream.completeFile(DFSOutputStream.java:2160)
at org.apache.hadoop.hdfs.DFSOutputStream.close(DFSOutputStream.java:2128)
at org.apache.hadoop.fs.FSDataOutputStream$PositionCache.close(FSDataOutputStream.java:70)
at org.apache.hadoop.fs.FSDataOutputStream.close(FSDataOutputStream.java:103)
at com.huawei.pai.collect2.stream.SequenceFileInfoChannel.close(SequenceFileInfoChannel.java:433)
at com.huawei.pai.collect2.stream.SequenceFileWriterToolChannel
```

```
$FileCloseTask.call(SequenceFileWriterToolChannel.java:804)
at com.huawei.pai.collect2.stream.SequenceFileWriterToolChannel
$FileCloseTask.call(SequenceFileWriterToolChannel.java:792)
at java.util.concurrent.FutureTask.run(FutureTask.java:262)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1145)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:615)
at java.lang.Thread.run(Thread.java:745)
```

## Cause Analysis

1. The HDFS client starts to write blocks.

For example, the HDFS client starts to write /  
**20150527/10/6\_20150527105000\_20150527105500\_SR5S14\_143272380633  
8\_128\_11.pkg.tmp1432723806338 at 2015-05-27 18:50:24,232**. The  
allocated block is **blk\_1099105501\_25370893**:

```
2015-05-27 18:50:24,232 | INFO | IPC Server handler 30 on 25000 | BLOCK* allocateBlock: /
20150527/10/6_20150527105000_20150527105500_SR5S14_1432723806338_128_11.pkg.tmp1432723
806338_BP-1803470917-192.168.57.33-1428597734132
blk_1099105501_25370893{blockUCState=UNDER_CONSTRUCTION, primaryNodeIndex=-1,
replicas=[ReplicaUnderConstruction[[DISK]DS-b2d7b7d0-f410-4958-8eba-6deecbca2f87:NORMAL|
RBW], ReplicaUnderConstruction[[DISK]DS-76bd80e7-ad58-49c6-bf2c-03f91caf750f:NORMAL|RBW]]}
|
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.saveAllocatedBlock(FSNamesystem.java:3166
)
```

2. After the writing is complete, the HDFS client invokes **fsync**:

```
2015-05-27 19:00:22,717 | INFO | IPC Server handler 22 on 25000 | BLOCK* fsync:
20150527/10/6_20150527105000_20150527105500_SR5S14_1432723806338_128_11.pkg.tmp1432723
806338 for DFSClient_NONMAPREDUCE_-120525246_15 |
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.fsync(FSNamesystem.java:3805)
```

3. The HDFS client invokes **close** to close the file. After receiving the close request from the client, the NameNode uses the `checkFileProgress` function to check the completion status of the last block and closes the file only when enough DataNodes report that the last block is complete:

```
2015-05-27 19:00:27,603 | INFO | IPC Server handler 44 on 25000 | BLOCK* checkFileProgress:
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}
has not reached minimal replication 1 |
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkFileProgress(FSNamesystem.java:3197)
2015-05-27 19:00:28,005 | INFO | IPC Server handler 45 on 25000 | BLOCK* checkFileProgress:
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}
has not reached minimal replication 1 |
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkFileProgress(FSNamesystem.java:3197)
2015-05-27 19:00:28,806 | INFO | IPC Server handler 63 on 25000 | BLOCK* checkFileProgress:
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}
has not reached minimal replication 1 |
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkFileProgress(FSNamesystem.java:3197)
2015-05-27 19:00:30,408 | INFO | IPC Server handler 43 on 25000 | BLOCK* checkFileProgress:
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}
has not reached minimal replication 1 |
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkFileProgress(FSNamesystem.java:3197)
2015-05-27 19:00:33,610 | INFO | IPC Server handler 37 on 25000 | BLOCK* checkFileProgress:
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}
has not reached minimal replication 1 |
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkFileProgress(FSNamesystem.java:3197)
2015-05-27 19:00:40,011 | INFO | IPC Server handler 37 on 25000 | BLOCK* checkFileProgress:
```

```
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,  
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|  
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}  
has not reached minimal replication 1 |  
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkFileProgress(FSNamesystem.java:3197)
```

- The NameNode prints **CheckFileProgress** multiple times because the HDFS client retries to close the file for several times. The file closing fails because the block status is not complete. The number of retries is determined by the **dfs.client.block.write.locateFollowingBlock.retries** parameter. The default value is 5. Therefore, **CheckFileProgress** is printed six times in the NameNode log.
- After 0.5 seconds, the DataNodes report that the block has been successfully written.

```
2015-05-27 19:00:40,608 | INFO | IPC Server handler 60 on 25000 | BLOCK* addStoredBlock:  
blockMap updated: 192.168.10.21:25009 is added to  
blk_1099105501_25370893{blockUCState=COMMITTED, primaryNodeIndex=-1,  
replicas=[ReplicaUnderConstruction[[DISK]DS-ef5fd3c9-5088-4813-ae9a-34a0714ec3a3:NORMAL|  
RBW], ReplicaUnderConstruction[[DISK]DS-f863e30f-ce5b-48cc-9cca-72f64c558adc:NORMAL|RBW]]}  
size 11837530 |  
org.apache.hadoop.hdfs.server.blockmanagement.BlockManager.logAddStoredBlock(BlockManager.java  
:2393)  
2015-05-27 19:00:48,297 | INFO | IPC Server handler 37 on 25000 | BLOCK* addStoredBlock:  
blockMap updated: 192.168.10.10:25009 is added to blk_1099105501_25370893 size 11837530 |  
org.apache.hadoop.hdfs.server.blockmanagement.BlockManager.logAddStoredBlock(BlockManager.java  
:2393)
```
- The block write success notification is delayed because of network bottlenecks or CPU bottlenecks.
- If close is invoked again or the number of file closing retries increases, a closing success message will be displayed. You are advised to increase the value of **dfs.client.block.write.locateFollowingBlock.retries**. The default parameter value is 5 and retry intervals are 400 ms, 800 ms, 1600 ms, 3200 ms, 6400 ms, and 12800 ms. Therefore, the result of the close function can be returned after a maximum of 25.2 seconds.

## Solution

Set the value of **dfs.client.block.write.locateFollowingBlock.retries** to 6. The retry intervals are 400 ms, 800 ms, 1600 ms, 3200 ms, 6400 ms, and 12800 ms. Therefore, the result of the close function can be returned after a maximum of 50.8 seconds.

## Remarks

Generally, this fault occurs when the cluster workload is heavy. Adjusting the parameter can only temporarily avoid the fault. You are advised to reduce the cluster workload.

For example, do not allocate all CPU resources to MapReduce.

## 12.15 File Fails to Be Uploaded to HDFS Due to File Errors

### Symptom

The **hadoop dfs -put** command is used to copy local files to HDFS.

After some files are uploaded, an error occurs. The size of the temporary files no long changes on the native NameNode page.

## Cause Analysis

1. Check the NameNode log `/var/log/Bigdata/hdfs/nn/hadoop-omm-namenode-hostname.log`. It is found that the file is being written until a failure occurs.

```
2015-07-13 10:05:07,847 | WARN | org.apache.hadoop.hdfs.server.namenode.LeaseManager
$Monitor@36fea922 | DIR* NameSystem.internalReleaseLease: Failed to release lease for file /hive/
order/OS_ORDER_8.txt_COPYING_. Committed blocks are waiting to be minimally replicated. Try
again later. | FSNamesystem.java:3936
2015-07-13 10:05:07,847 | ERROR | org.apache.hadoop.hdfs.server.namenode.LeaseManager
$Monitor@36fea922 | Cannot release the path /hive/order/OS_ORDER_8.txt_COPYING_ in the lease
[Lease. Holder: DFSCClient_NONMAPREDUCE_-1872896146_1, pendingcreates: 1] |
LeaseManager.java:459
org.apache.hadoop.hdfs.protocol.AlreadyBeingCreatedException: DIR*
NameSystem.internalReleaseLease: Failed to release lease for file /hive/order/
OS_ORDER_8.txt_COPYING_. Committed blocks are waiting to be minimally replicated. Try again
later.
at FSNamesystem.internalReleaseLease(FSNamesystem.java:3937)
```
2. Root cause: The uploaded files are damaged.
3. Verification: The cp or scp operation fails to be performed for the copied files. Therefore, the files are damaged.

## Solution

**Step 1** Upload normal files.

----End

## 12.16 After `dfs.blocksize` Is Configured on the UI and Data Is Uploaded, the Block Size Does Not Change

### Symptom

After `dfs.blocksize` is set to `268435456` on the UI and data is uploaded, the original block size keeps unchanged.

### Cause Analysis

The `dfs.blocksize` value in the `hdfs-site.xml` file of the HDFS client is not changed, and this value is used.

### Solution

**Step 1** Ensure that the `dfs.blocksize` value is a multiple of 512.

**Step 2** Download a client or modify the client configuration.

**Step 3** `dfs.blocksize` is configured on the client and is subject to the client. Otherwise, the value configured on the server is used.

----End



## 12.17 HDFS File Fails to Be Read, and Error Message "FileNotFoundException" Is Displayed

### Symptom

In MapReduce tasks, all Map tasks are successfully executed, but Reduce tasks fail. The error message "FileNotFoundException...No lease on...File does not exist" is displayed in the logs.

```
Error: org.apache.hadoop.ipc.RemoteException(java.io.FileNotFoundException): No lease on /user/sparkhive/warehouse/daas/dsp/output/_temporary/1/_temporary/attempt_1479799053892_17075_r_000007_0/part-r-00007 (inode 6501287): File does not exist. Holder DFSClient_attempt_1479799053892_17075_r_000007_0_-1463597952_1 does not have any open files.
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkLease(FSNamesystem.java:3350)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.completeFileInternal(FSNamesystem.java:3442)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.completeFile(FSNamesystem.java:3409)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.complete(NameNodeRpcServer.java:789)
```

### Cause Analysis

"FileNotFoundException...No lease on...File does not exist" indicates that the file is deleted during the operation.

1. Search for the file name in the NameNode audit log of HDFS (**/var/log/Bigdata/audit/hdfs/nn/hdfs-audit-namenode.log** of the active NameNode) to confirm the creation time of the file.
2. Search the NameNode audit logs that are generated within the time range from the file creation to the time of exception occurrence and determine whether the file is deleted or moved to another directory.
3. If the file is not deleted or moved, the parent directory of the file may be deleted or moved. You need to search the upper-layer directory. In this example, the parent directory of the file is deleted.

```
2017-05-31 02:04:08,286 | INFO | IPC Server handler 30 on 25000 | allowed=true
ugi=appUser@HADOOP.COM (auth:TOKEN) ip=/192.168.1.22 cmd=delete src=/user/sparkhive/warehouse/daas/dsp/output/_temporary dst=null perm=null proto=rpc | FSNamesystem.java:8189
```

#### NOTE

- The preceding log indicates that the **appUser** user of the 192.168.1.22 node deletes **/user/sparkhive/warehouse/daas/dsp/output/\_temporary**.
- Run the **zgrep "file name" \*.zip** command to search for the contents of the .zip package.

### Solution

- Step 1** Check the service to find out why the file or the parent directory of the file is deleted.

----End

## 12.18 Failed to Write Files to HDFS, and Error Message "item limit of xxx is exceeded" Is Displayed

### Symptom

The client or upper-layer component logs indicate that a file fails to be written to a directory on HDFS. The error information is as follows:

The directory item limit of /tmp is exceeded: limit=5 items=5.

### Cause Analysis

1. The run log file `/var/log/Bigdata/hdfs/nn/hadoop-omm-namenode-XXX.log` of the client or NameNode contains error information "The directory item limit of /tmp is exceeded:." The error message indicates that the number of files in the `/tmp` directory exceeds 1048576.  

```
2018-03-14 11:18:21,625 | WARN | IPC Server handler 62 on 25000 | DIR* NameSystem.startFile: /tmp/test.txt The directory item limit of /tmp is exceeded: limit=1048576 items=1048577 | FSNamesystem.java:2334
```
2. The `dfs.namenode.fs-limits.max-directory-items` parameter specifies the maximum number of directories or files that are not in recursion relationship in a single directory. The default value is **1048576**. The value ranges from 1 to 6400000.

### Solution

**Step 1** Check whether it is normal that the directory contains more than one million files that are not in recursion relationship. If it is normal, increase the value of the HDFS parameter `dfs.namenode.fs-limits.max-directory-items` and restart the HDFS NameNode for the modification to take effect.

**Step 2** If it is abnormal, delete unnecessary files.

----End

## 12.19 Adjusting the Log Level of the HDFS SHDFS Shell Client

- **Temporary adjustment:** After the Shell client window is closed, the log is restored to the default value.
  - a. Run the `export HADOOP_ROOT_LOGGER` command to adjust the log level of the client.
  - b. Run the `export HADOOP_ROOT_LOGGER=Log level,console` command to adjust the log level of the shell client.  
Run the `export HADOOP_ROOT_LOGGER=DEBUG,console` command to adjust the log level to **Debug**.  
Run the `export HADOOP_ROOT_LOGGER=ERROR,console` command to adjust the log level to **Error**.

- **Permanent adjustment**
  - a. Add **export HADOOP\_ROOT\_LOGGER=Log level,console** to the HDFS client's environment variable configuration file **/opt/client/HDFS/component\_env** (replace **/opt/client** with the actual client path).
  - b. Run the **source /opt/client/bigdata\_env** command.
  - c. Run the command on the client again.

## 12.20 HDFS File Read Fails, and Error Message "No common protection layer" Is Displayed

### Symptom

HDFS fails to be operated on the Shell client or other clients, and the error message "No common protection layer between client and server" is displayed.

Running any **hadoop** command, such as **hadoop fs -ls /**, on a node outside the cluster fails. The bottom-layer error message is displayed stating "No common protection layer between client and server."

```
2017-05-13 19:14:19,060 | ERROR | [pool-1-thread-1] | Server startup failure |
org.apache.sqoop.core.SqoopServer.initializeServer(SqoopServer.java:69)
org.apache.sqoop.common.SqoopException: MAPRED_EXEC_0028:Failed to operate HDFS - Failed to get the
file /user/loader/etl_dirty_data_dir status
    at org.apache.sqoop.job.mr.HDFSClient.fileExist(HDFSClient.java:85)
...
    at java.lang.Thread.run(Thread.java:745)
Caused by: java.io.IOException: Failed on local exception: java.io.IOException: Couldn't setup connection for
loader/hadoop@HADOOP.COM to loader37/10.162.0.37:25000; Host Details : local host is:
"loader37/10.162.0.37"; destination host is: "loader37":25000;
    at org.apache.hadoop.net.NetUtils.wrapException(NetUtils.java:776)
...
... 10 more
Caused by: java.io.IOException: Couldn't setup connection for loader/hadoop@HADOOP.COM to
loader37/10.162.0.37:25000
    at org.apache.hadoop.ipc.Client$Connection$1.run(Client.java:674)
... 28 more
Caused by: javax.security.sasl.SaslException: No common protection layer between client and server
    at com.sun.security.sasl.gsskerb.GssKrb5Client.doFinalHandshake(GssKrb5Client.java:251)
...
    at org.apache.hadoop.ipc.Client$Connection.setupIOStreams(Client.java:720)
```

### Cause Analysis

1. The RPC protocol is used for data transmission between the client and server of HDFS. The protocol has multiple encryption modes and the **hadoop.rpc.protection** parameter specifies the mode to use.
2. If the value of the **hadoop.rpc.protection** parameter on the client is different from that on the server, the "No common protection layer between client and server" error is reported.

 NOTE

**hadoop.rpc.protection** indicates that data can be transmitted between nodes in any of the following modes:

- **privacy**: Data is transmitted after authentication and encryption. This mode reduces the performance.
- **authentication**: Data is transmitted after authentication without encryption. This mode ensures performance but has security risks.
- **integrity**: Data is transmitted without encryption or authentication. To ensure data security, exercise caution when using this mode.

## Solution

**Step 1** Download the client again. If the client is an application, update the configuration file in the application.

----End

## 12.21 Failed to Write Files Because the HDFS Directory Quota Is Insufficient

### Symptom

After the quota is set for a directory, writing files to the directory fails. Error message "The DiskSpace quota of /tmp/tquota2 is exceeded" is displayed.

```
[omm@189-39-150-115 client]$ hdfs dfs -put switchuser.py /tmp/tquota2  
put: The DiskSpace quota of /tmp/tquota2 is exceeded: quota = 157286400 B = 150 MB but disk space  
consumed = 402653184 B = 384 MB
```

### Possible Causes

The remaining space configured for the directory is less than the space required for writing files.

### Cause Analysis

1. HDFS supports setting the quota for a specific directory, that is, the maximum space occupied by files in a directory can be set. For example, the following command is used to set a maximum of 150 MB files to be written to the **/tmp/tquota** directory (Space = Block size x Number of replicas):

```
hadoop dfsadmin -setSpaceQuota 150M /tmp/tquota2
```

2. Run the following command to check the configured quota for the directory. **SPACE\_QUOTA** is the configured space quota, and **REM\_SPACE\_QUOTA** is the remaining space.

```
hdfs dfs -count -q -h -v /tmp/tquota2
```

**Figure 12-11** Viewing the quota set for a directory

```
hdfs dfs -count -q -h -v /tmp/tquota2  
QUOTA REM_QUOTA SPACE_QUOTA REM_SPACE_QUOTA DIR_COUNT FILE_COUNT CONTENT_SIZE PATHNAME  
none inf 150M 150M 1 0 0 /tmp/tquota2
```

3. Analyze logs. The following log indicates that writing the file requires 384 MB space, but the current space quota is only 150 MB. Therefore, the space is insufficient. Before a file is written, the required remaining space is as follows: Block size x Number of copies. 128 MB x 3 copies = 384 MB.

```
[omm@189-39-150-115 client]$  
[omm@189-39-150-115 client]$ hdfs dfs -put switchuser.py /tmp/tquota2  
put: The DiskSpace quota of /tmp/tquota2 is exceeded: quota = 157286400 B = 150 MB but diskpace  
consumed = 402653184 B = 384 MB
```

## Solution

- Step 1** Set a proper quota for the directory.

```
hadoop dfsadmin -setSpaceQuota 150G /directory name
```

- Step 2** Run the following command to clear the quota:

```
hdfs dfsadmin -clrSpaceQuota /directory name
```

----End

## 12.22 Balancing Fails, and Error Message "Source and target differ in block-size" Is Displayed

### Symptom

When the **distcp** command is executed to copy files across clusters, the message "Source and target differ in block-size." is displayed, indicating that some files fail to be copied. Use **-pb** to preserve block-sizes during copy. "

```
Caused by: java.io.IOException: Check-sum mismatch between hdfs://10.180.144.7:25000/kylin/  
kylin_default_instance_prod/parquet/f2e72874-f01c-45ff-b219-207f3a5b3fcb/c769cd2d-575a-4459-837b-  
a19dd7b20c27/339114721280/0.parquet and hdfs://10.180.180.194:25000/kylin/  
kylin_default_instance_prod/parquet/f2e72874-f01c-45ff-  
b219-207f3a5b3fcb/.distcp.tmp.attempt_1523424430246_0004_m_000019_2. Source and target differ in  
block-size. Use -pb to preserve block-sizes during copy. Alternatively, skip checksum-checks altogether,  
using -skipCrc. (NOTE: By skipping checksums, one runs the risk of masking data-corruption during file-  
transfer.) at  
org.apache.hadoop.tools.mapred.RetriableFileCopyCommand.compareCheckSums(RetriableFileCopyComman  
d.java:214)
```

### Possible Causes

This is not a version-related problem. When you run the **distcp** command to copy files, the block size of the source file is not recorded by default. As a result, the verification fails when the block size of the source file is not 128 MB. In this case, you need to add parameter **-pb** to the **distcp** command.

### Cause Analysis

1. The block size is set when data is written to HDFS. The default block size is 128 MB. The size of files written by some components or service programs may not be 128 MB, for example, 8 MB.

```
<name>dfs.blocksize</name>  
<value>134217728</value>
```

**Figure 12-12** Size of files written by some components or service programs

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rwxrwx----	bill	hive	17.9 MB	Wed Dec 13 17:22:44 2017	3	8 MB	/user/hive/warehouse/orctest.db/new_orc_07/enddate=20171202/part-00000

- DistCp reads the file from a source cluster and writes it to a destination cluster. By default, the value of `dfs.blocksize` in the MapReduce task is used as the block size, whose default value is 128 MB.
- After DistCp finishes writing a file, the system performs verification based on the physical size of the block. Because the block size of the file in the source cluster is different from that of the file in the destination cluster, the splitting sizes are different. As a result, the verification fails.

In the preceding file, there are three blocks ( $17.9/8 \text{ MB} = 3$ ) in the old cluster and one block ( $17.9/128 \text{ MB} = 1$ ) in the new cluster. As a result, the different physical block sizes cause the verification failure.

## Solution

Add parameter `-pb` in the `distcp` command. This parameter is used to reserve the block size when `distcp` is used to ensure that the block size of the new cluster is the same as that of the old cluster.

**Figure 12-13** Size of the reserved block during `distcp` command execution

```
[root@189-39-235-118 clientu10]# hadoop distcp -pb hdfs://haclusterX/user hdfs://hacluster/tmp/test
```

## 12.23 Failed to Query or Delete HDFS Files

### Symptom

A file fails to be queried or deleted using the HDFS shell client. The file (invisible characters) can be viewed in the parent directory.

**Figure 12-14** List of files in the parent directory

```
drwxrwx---- - datalab90020_639_w hive 0 2018-04-10 01:44 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp
drwxrwx---- - datalab90020_639_w hive 0 2018-04-10 16:45 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp2
[root@dgtsp355-or-FusionInsight_client]# hadoop fs -ls /user/hive/warehouse/datalake_dwi_barpsit.db
Found 4 items
drwxrwxr-x - datalab90020_639_w hive 0 2018-04-11 12:05 /user/hive/warehouse/datalake_dwi_barpsit.db/bak_v_tp_mp_aut_input
drwxrwx---- - datalab90020_639_w hive 0 2018-04-11 11:16 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
drwxrwx---- - datalab90020_639_w hive 0 2018-04-10 01:44 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp
drwxrwx---- - datalab90020_639_w hive 0 2018-04-10 16:45 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp2
[root@dgtsp355-or-FusionInsight_client]# hadoop fs -rm -r /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
rm: /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input: No such file or directory
[root@dgtsp355-or-FusionInsight_client]# hdfs dfs -rm -r /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
rm: /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input: No such file or directory
[root@dgtsp355-or-FusionInsight_client]#
[root@dgtsp355-or-FusionInsight_client]# hdfs dfs -ls /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
ls: /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input: No such file or directory
[root@dgtsp355-or-FusionInsight_client]#
[root@dgtsp355-or-FusionInsight_client]#
```

### Cause Analysis

The possible cause is that invisible characters are written to the file. You can write the file name to the local text and run the `vi` command to open the file.

```
hdfs dfs -ls parent directory > /tmp/t.txt
```

```
vi /tmp/t.txt
```

Run the **:set list** command to display invisible characters in the file name. For example, the file name contains **^M**, which is invisible.

Figure 12-15 Displaying invisible characters

```
Found 1 items
drwxrwx--- - datalab90020_639_w hive 0 2018-04-11 11:16 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input^M
```

## Solution

- Step 1** Run the Shell command to read the file name recorded in the text. Ensure that the following command output contains the full path of the file in HDFS.

```
cat /tmp/t.txt |awk '{print $8}'
```

Figure 12-16 File path

```
drwxrwx--- - datalab90020_639_w hive 0 2018-04-11 11:16 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
drwxrwx--- - datalab90020_639_w hive 0 2018-04-10 01:44 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp
drwxrwx--- - datalab90020_639_w hive 0 2018-04-10 16:45 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp2
[root@dggts355-or-FusionInsight_client]# cat /tmp/t.txt |awk '{print $8}'
/user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
[root@dggts355-or-FusionInsight_client]# hadoop fs -rm -r $(cat /tmp/t.txt |awk '{print $8}')
to trash at: hdfs://hacluster/user/datalab90020_639_w/.Trash/Current/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input
[root@dggts355-or-FusionInsight_client]# hdfs dfs -ls /user/hive/warehouse/datalake_dwi_barpsit.db
Found 2 items
drwxrwx--- - datalab90020_639_w hive 0 2018-04-10 01:44 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp
drwxrwx--- - datalab90020_639_w hive 0 2018-04-10 16:45 /user/hive/warehouse/datalake_dwi_barpsit.db/v_tp_mp_aut_input_tmp2
[root@dggts355-or-FusionInsight_client]#
```

- Step 2** Run the following command to delete the file:

```
hdfs dfs -rm $(cat /tmp/t.txt |awk '{print $8}')
```

- Step 3** Verify that the file has been deleted.

```
hdfs dfs -ls Parent directory
```

----End

## 12.24 Uneven Data Distribution Due to Non-HDFS Data Residuals

### Symptom

Data distribution is uneven. A disk is full while other disks have sufficient space.

The data storage directory of HDFS DataNode is set to **/export/data1/dfs--/export/data12/dfs**. A large volume of data is stored to **/export/data1/dfs** but data is evenly distributed to other disks.

### Cause Analysis

The customer's disk is reinstalled. However, a directory is not thoroughly deleted during disk uninstallation, that is, the added disk is unformatted and historical junk data remains.

### Solution

Manually delete data residuals.

## 12.25 Uneven Data Distribution Due to HDFS Client Installation on the DataNode

### Symptom

Data is unevenly distributed on HDFS DataNodes. Disk usage of a node is high or even reaches 100% while disks on other nodes have sufficient idle space.

### Cause Analysis

In the HDFS data replica mechanism, the first replica is stored to the local node where the client is stored. As a result, disks of the node run out while disks of other nodes have sufficient idle space.

### Solution

**Step 1** For the existing data unevenly distributed, run the following command to balance data:

```
/opt/client/HDFS/hadoop/sbin/start-balancer.sh -threshold 10
```

**/opt/client** indicates the actual client installation directory.

**Step 2** For new data, install the client on the node without DataNode.

----End

## 12.26 Unbalanced DataNode Disk Usages of a Node

### Symptom

The disk usage of each DataNode on a node is uneven.

Example:

```
189-39-235-71:~ # df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/xvda       360G  92G   250G  28% /
/dev/xvdb       700G  900G   200G  78% /srv/BigData/hadoop/data1
/dev/xvdc       700G  900G   200G  78% /srv/BigData/hadoop/data2
/dev/xvdd       700G  900G   200G  78% /srv/BigData/hadoop/data3
/dev/xvde       700G  900G   200G  78% /srv/BigData/hadoop/data4
/dev/xvdf       10G   900G  890G   2% /srv/BigData/hadoop/data5
189-39-235-71:~ #
```

### Possible Causes

Some disks are faulty and are replaced with new ones. The new disk usage is low.

Disks are added. For example, the original four data disks are expanded to five disks.



## Cause Analysis

There are two policies for writing data to Block disks on DataNodes: 1. Round Robin (default value) and 2. Preferentially writing data to the disk with the more available space.

Description of the **dfs.datanode.fsdataset.volume.choosing.policy** parameter

Possible values:

- Polling:  
**org.apache.hadoop.hdfs.server.datanode.fsdataset.RoundRobinVolumeChoosingPolicy**
- Preferentially writing data to the disk with more available space:  
**org.apache.hadoop.hdfs.server.datanode.fsdataset.AvailableSpaceVolumeChoosingPolicy**

## Solution

Change the value of **dfs.datanode.fsdataset.volume.choosing.policy** to **org.apache.hadoop.hdfs.server.datanode.fsdataset.AvailableSpaceVolumeChoosingPolicy**, save the settings, and restart the affected services or instances.

In this way, the DataNode preferentially selects a node with the most available disk space to store data copies.

### NOTE

- Data written to the DataNode will be preferentially written to the disk with more available disk space.
- The high usage of some disks can be relieved with the gradual deletion of aging data from the HDFS.

## 12.27 Locating Common Balance Problems

### Problem 1: Lack of Permission to Execute the Balance Task (Access denied)

Problem details: After the **start-balancer.sh** command is executed, the "hadoop-root-balancer-Hostname.out" log displays "Access denied for user test1. Superuser privilege is required."

```
cat /opt/client/HDFS/hadoop/logs/hadoop-root-balancer-host2.out
Time Stamp      Iteration#  Bytes Already Moved  Bytes Left To Move  Bytes Being Moved
INFO: Watching file:/opt/client/HDFS/hadoop/etc/hadoop/log4j.properties for changes with interval : 60000
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.AccessControlException): Access denied
for user test1.
Superuser privilege is required
at
org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSuperuserPrivilege(FSPermissionChecker
.java:122)
at
org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkSuperuserPrivilege(FSNamesystem.java:5916)
```

#### Cause analysis:

The administrator account is required for executing the balance task.

#### Solution:

- Secure version  
Perform authentication for user **hdfs** or a user in the **supergroup** group and then execute the balance task.
- Normal version  
Run the **su - hdfs** command on the client before running the **balance** command on HDFS.

## Problem 2: Failed to Execute the Balance Task, and /system/balancer.id Reports an Exception

### Problem details:

A user starts a balance process on the HDFS client. After the process is stopped unexpectedly, the user performs the balance operation again. The operation fails.

```
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.protocol.RecoveryInProgressException): Failed to APPEND_FILE /system/balancer.id for DFSClient because lease recovery is in progress. Try again later.
```

### Cause analysis:

Generally, after the balance operation is complete in HDFS, the **/system/balancer.id** file is automatically released and the balance operation can be performed again.

In the preceding scenario, the first balance operation is stopped abnormally. Therefore, when the balance operation is performed for the second time, the **/system/balancer.id** file still exists. As a result, the **append /system/balancer.id** operation is triggered and the balance operation fails.

### Solution:

Method 1: After the hard lease period exceeds one hour, release the lease on the original client and perform the balance operation again.

Method 2: Delete the **/system/balancer.id** file from HDFS and perform the balance operation again.

## 12.28 HDFS Displays Insufficient Disk Space But 10% Disk Space Remains

### Symptom

1. The alarm "HDFS Disk Usage Exceeds the Threshold" is reported.
2. On the HDFS page, high disk space usage is displayed.

### Cause Analysis

The **dfs.datanode.du.reserved.percentage** parameter is set in HDFS, indicating the percentage of the reserved space of each disk to the total disk space. The DataNode reserves space you set for NodeManager running and computing of other components, for example, Yarn, or for upgrades.

As 10% disk space is reserved, the HDFS DataNode regards that there is no available disk space when the disk usage reaches 90%.

## Solution

- Step 1** Expand the capacity when the HDFS DataNode disk usage reaches 80%. For details, see [Scaling Out a Cluster](#). After the scale-out is complete, balance data. For details, see [Balancing DataNode Capacity](#).
- Step 2** If the disk capacity cannot be expanded in time, delete useless data in HDFS to release disk space.
- End

## 12.29 Error Message "error creating DomainSocket" Is Displayed When the HDFS Client Installed on the Core Node in a Normal Cluster Is Used

### Issue

In a normal cluster, an error message is displayed when a user is created on the core node to install the HDFS client.

### Symptom

In a common cluster, the following error message is displayed when a user is created on the Core node to install the client:

```
2020-03-14 19:16:17,166 WARN shortcircuit.DomainSocketFactory: error creating DomainSocket
java.net.ConnectException: connect(2) error: Permission denied when trying to connect to '/var/run/MRS-
HDFS/dn_socket'
at org.apache.hadoop.net.unix.DomainSocket.connect0(Native Method)
at org.apache.hadoop.net.unix.DomainSocket.connect(DomainSocket.java:256)
at org.apache.hadoop.hdfs.shortcircuit.DomainSocketFactory.createSocket(DomainSocketFactory.java:168)
at org.apache.hadoop.hdfs.client.impl.BlockReaderFactory.nextDomainPeer(BlockReaderFactory.java:799)
...
```

### Cause Analysis

A user runs the **useradd** command to create a user. The default user group of the user does not contain the **ficommon** user group. As a result, the preceding error is reported when the **get** command of HDFS is executed.

### Procedure

Run the **usermod -a -G ficommon username** command to add the user to the **ficommon** user group.

## 12.30 HDFS Files Fail to Be Uploaded When the Client Is Installed on a Node Outside the Cluster

### Issue

A client installed on a node outside the cluster fails to upload files using **hdfs**.

## Symptom

After a client is installed on a cluster node and a file is uploaded using the **hdfs** command, the following error is reported.

Figure 12-17 Error reported during file upload

```
[root@ywwa02 bin]# hadoop fs -put test.txt /tmp/input
2020-07-31 18:12:27,533 INFO obs.OBSFileSystem: This FileSystem GC-ful, clear resource.
2020-07-31 18:12:31,757 INFO hdfs.DataStreamer: Exception in createBlockOutputStream blk_1073774851_34031
java.net.NoRouteToHostException: No route to host
    at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:717)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DataStreamer.createSocketForPipeline(DataStreamer.java:255)
    at org.apache.hadoop.hdfs.DataStreamer.createBlockOutputStream(DataStreamer.java:1789)
    at org.apache.hadoop.hdfs.DataStreamer.nextBlockOutputStream(DataStreamer.java:1743)
    at org.apache.hadoop.hdfs.DataStreamer.run(DataStreamer.java:718)
2020-07-31 18:12:31,759 WARN hdfs.DataStreamer: Abandoning BP-1721849101-192.168.0.86-1595473704426:blk_1073774851_34031
2020-07-31 18:12:31,800 WARN hdfs.DataStreamer: Excluding datanode DataNodeInfoWithStorage[192.168.0.157:9866,DS-592b7049-b4af-4bba-a184-1e1928a9028b,D1SK]
2020-07-31 18:12:34,869 INFO hdfs.DataStreamer: Exception in createBlockOutputStream blk_1073774852_34032
java.net.NoRouteToHostException: No route to host
    at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:717)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DataStreamer.createSocketForPipeline(DataStreamer.java:255)
    at org.apache.hadoop.hdfs.DataStreamer.createBlockOutputStream(DataStreamer.java:1789)
    at org.apache.hadoop.hdfs.DataStreamer.nextBlockOutputStream(DataStreamer.java:1743)
    at org.apache.hadoop.hdfs.DataStreamer.run(DataStreamer.java:718)
2020-07-31 18:12:34,869 WARN hdfs.DataStreamer: Abandoning BP-1721849101-192.168.0.86-1595473704426:blk_1073774852_34032
2020-07-31 18:12:34,899 WARN hdfs.DataStreamer: Excluding datanode DataNodeInfoWithStorage[192.168.0.189:9866,DS-5bee1b3a-4453-4d86-a632-262cb67c0bdb,D1SK]
2020-07-31 18:12:37,948 INFO hdfs.DataStreamer: Exception in createBlockOutputStream blk_1073774853_34033
java.net.NoRouteToHostException: No route to host
    at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:717)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DataStreamer.createSocketForPipeline(DataStreamer.java:255)
    at org.apache.hadoop.hdfs.DataStreamer.createBlockOutputStream(DataStreamer.java:1789)
    at org.apache.hadoop.hdfs.DataStreamer.nextBlockOutputStream(DataStreamer.java:1743)
    at org.apache.hadoop.hdfs.DataStreamer.run(DataStreamer.java:718)
2020-07-31 18:12:37,948 WARN hdfs.DataStreamer: Abandoning BP-1721849101-192.168.0.86-1595473704426:blk_1073774853_34033
2020-07-31 18:12:37,988 WARN hdfs.DataStreamer: Excluding datanode DataNodeInfoWithStorage[192.168.0.174:9866,DS-fa34f00b-2c03-4d0e-ad5e-3a2555735c0b,D1SK]
2020-07-31 18:12:38,034 WARN hdfs.DataStreamer: DataStreamer Exception
org.apache.hadoop.ipc.RemoteException(java.io.IOException): File /tmp/input/test.txt_COPYING could only be written to 0 of the 1 minReplication nodes. There are 3 da
    at org.apache.hadoop.hdfs.server.blockmanagement.BlockManager.chooseTarget4NewBlock(BlockManager.java:2223)
    at org.apache.hadoop.hdfs.server.namenode.FSDataWriterImpl.chooseTargetForNewBlock(FSDataWriterImpl.java:346)
    at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getAdditionalBlock(FSNamesystem.java:2727)
    at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.addBlock(NameNodeRpcServer.java:879)
    at org.apache.hadoop.hdfs.protocolPB.ClientNameNodeProtocolServerSideTranslatorPB.addBlock(ClientNameNodeProtocolServerSideTranslatorPB.java:596)
    at org.apache.hadoop.hdfs.protocol.proto.ClientNameNodeProtocolProtos$ClientNameNodeProtocol$2.callBlockingMethod(ClientNameNodeProtocolProtos.java)
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:530)
    at org.apache.hadoop.ipc.RPCServer.call(RPC.java:1036)
```

## Cause Analysis

The error message "no route to host" is displayed, and the IP address 192.168 is contained in the error message. That is, the internal network route from the client node to the DataNode in the cluster is unreachable. As a result, the file fails to be uploaded.

## Procedure

In the client directory of the client node, find the **hdfs-site.xml** file in the HDFS client configuration directory. Add the **dfs.client.use.datanode.hostname** configuration item to the configuration file, and set the value to **true**.

## 12.31 Insufficient Number of Replicas Is Reported During High Concurrent HDFS Writes

### Symptom

File writes to HDFS fail occasionally.

The operation log is as follows:

```
105 | INFO | IPC Server handler 23 on 25000 | IPC Server handler 23 on 25000, call
org.apache.hadoop.hdfs.protocol.ClientProtocol.addBlock from 192.168.1.96:47728 Call#1461167 Retry#0 |
Server.java:2278
```





## 12.33 An Error Is Reported When a Node Outside the Cluster Accesses MRS HDFS

### Symptom

When a node outside the MRS cluster accesses HDFS in the cluster, the following error message is reported: Class `org.apache.hadoop.hdfs.server.namenode.ha.AdaptiveFailoverProxyProvider` not found

```
java.lang.RuntimeException: java.lang.ClassNotFoundException: Class org.apache.hadoop.hdfs.server.namenode.ha.AdaptiveFailoverProxyProvider not found
    at org.apache.hadoop.conf.Configuration.getClass(Configuration.java:2696)
    at org.apache.hadoop.hdfs.NameNodeProxiesClient.getFailoverProxyProviderClass(NameNodeProxiesClient.java:206)
    at org.apache.hadoop.hdfs.NameNodeProxiesClient.createFailoverProxyProvider(NameNodeProxiesClient.java:237)
    at org.apache.hadoop.hdfs.NameNodeProxiesClient.createFailoverProxyProvider(NameNodeProxiesClient.java:225)
    at org.apache.hadoop.hdfs.NameNodeProxiesClient.createProxyClientProtocol(NameNodeProxiesClient.java:135)
    at org.apache.hadoop.hdfs.DFSClient.<init>(DFSClient.java:358)
    at org.apache.hadoop.hdfs.DFSClient.<init>(DFSClient.java:295)
    at org.apache.hadoop.hdfs.DistributedFileSystem.initializeDistributedFileSystem(java:186)
    at org.apache.hadoop.fs.FileSystem.createFileSystem(FileSystem.java:3459)
    at org.apache.hadoop.fs.FileSystem.access$200(FileSystem.java:125)
    at org.apache.hadoop.fs.FileSystemCache.getInternalFileSystem(java:3512)
    at org.apache.hadoop.fs.FileSystemCache.get(FileSystem.java:3480)
    at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:490)
    at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:230)
    at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:474)
    at org.apache.hadoop.fs.Path.getFileSystem(Path.java:371)
    at org.apache.hadoop.fs.shell.PathData.expandPath(PathData.java:329)
    at org.apache.hadoop.fs.shell.Command.expandArgument(Command.java:249)
    at org.apache.hadoop.fs.shell.Command.expandArguments(Command.java:232)
    at org.apache.hadoop.fs.shell.FsCommand.processArguments(FsCommand.java:106)
    at org.apache.hadoop.fs.shell.Command.run(Command.java:176)
    at org.apache.hadoop.fs.PShell.run(PShell.java:244)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
    at org.apache.hadoop.fs.PShell.main(PShell.java:411)
Caused by: java.lang.RuntimeException: java.lang.ClassNotFoundException: Class org.apache.hadoop.hdfs.server.namenode.ha.AdaptiveFailoverProxyProvider not found
    at org.apache.hadoop.conf.Configuration.getClass(Configuration.java:2664)
    at org.apache.hadoop.conf.Configuration.getClass(Configuration.java:2688)
    ... 24 more
Caused by: java.lang.ClassNotFoundException: Class org.apache.hadoop.hdfs.server.namenode.ha.AdaptiveFailoverProxyProvider not found
    at org.apache.hadoop.conf.Configuration.getClassByName(Configuration.java:2508)
    at org.apache.hadoop.conf.Configuration.getClass(Configuration.java:2662)
    ... 25 more
```

### Cause Analysis

The possible causes are as follows:

- An error is reported when an open-source HDFS client accesses HDFS of an MRS cluster.
- An error is reported when the JAR package is used to connect to HDFS of the MRS cluster (including connection to HDFS during task submission).

### Procedure

Method 1:

**Step 1** Locate the HDFS configuration file `hdfs-site.xml` used by the command or JAR package.

**Step 2** Modify the `dfs.client.failover.proxy.provider.hacluster` configuration as follows:

```
<property>
<name>dfs.client.failover.proxy.provider.hacluster</name>
<value>org.apache.hadoop.hdfs.server.namenode.ha.ConfiguredFailoverProxyProvider</value>
</property>
```

#### NOTE

You can also delete the preceding configuration items.

**Step 3** Save the file and access MRS HDFS again.

----End

Method 2:

- Step 1** Download the hadoop-plugins matching the MRS cluster version from the Maven repository.

Download path: <https://repo.huaweicloud.com/repository/maven/huaweicloudsdk/com/huawei/mrs/hadoop-plugins/>

- Step 2** Add the downloaded JAR package to the dependency of the command or JAR package.

----End

## 12.34 "ALM-12027 Host PID Usage Exceeds the Threshold" Is Generated for a NameNode

### Symptom

In cluster 3.1.2 and earlier 3.x versions, alarm "ALM-12027 Host PID Usage Exceeds the Threshold" is generated for a NameNode, and Java processes of the node may report error "unable to create new native thread".

### Cause Analysis

1. Run the following command to collect statistics on the number of threads of node processes and sort the threads:

```
ps -eFT | awk '{print $2}' |sort -n |uniq -c |sort -n
```

The result is as follows.

```
64 15310
54533 2346
275 2983
504 32139
323 33641
113 38454
105 41630
174 43672
101 44918
66 55187
1 55315
76 7057
```

2. Check the process that starts the most threads. In this example, process 2346 is the NameNode process, which starts 54,000 threads and keeps increasing.
3. The jstack log of that process is printed for multiple times. According to the jstack log information, a large number of NameNode threads are in the WAITING state and have not been released for a long time.

```
Line 200: at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
Line 215: at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
Line 230: at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
Line 245: at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
Line 260: at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
Line 275: at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
```



```
"Thread-121224" #653415 daemon prio=5 os_prio=0 tid=0x000007fb25b753000 nid=0x50ea waiting on condition [0x00007fb1ea287000]
java.lang.Thread.State: WAITING (parking)
  at sun.misc.Unsafe.park(Native Method)
  - parking to wait for <0x000000073f5e8838> (a java.util.concurrent.locks.AbstractQueuedSynchronizer$ConditionObject)
  at java.util.concurrent.locks.LockSupport.park(LockSupport.java:175)
  at java.util.concurrent.locks.AbstractQueuedSynchronizer$ConditionObject.await(AbstractQueuedSynchronizer.java:2045)
  at java.util.concurrent.LinkedBlockingQueue.take(LinkedBlockingQueue.java:442)
  at java.util.concurrent.ThreadPoolExecutor.getTask(ThreadPoolExecutor.java:1074)
  at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1134)
  at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
  at java.lang.Thread.run(Thread.java:748)

Locked ownable synchronizers:
- None
```

According to the preceding analysis, the NameNode has a built-in mechanism to automatically enable the DEBUG log function based on the WARN log information. In the environment, the DEBUG log function is repeatedly enabled and log4j is continuously modified because the replica fails to be selected. After the log4j of the component is modified, the process automatically loads the configuration file, new threads are automatically generated, and this alarm is triggered after a long period of time.

In this case, disable the built-in mechanism to disable the function of automatically changing the log level.

## Procedure

- Step 1** Log in to the active and standby NameNodes in the cluster and run the following commands to back up the script:

```
cd $BIGDATA_HOME/FusionInsight_Current/*_*_NameNode/install/hadoop/
sbin
```

```
cp hdfs-namenode-period-check.sh /tmp
```

- Step 2** Edit the `hdfs-namenode-period-check.sh` file on the active and standby NameNodes.

```
vi hdfs-namenode-period-check.sh
```

Comment out `checkBlockplacementLog` in the `main` method. For example:

```
main()
{
  Log $INFO "start period check"
  checkHaState
  checkDefaultFS
  checkAutoBalancer
  checkFsMonitorDirectory
  checkAutoMover
  checkAutoDatamove
  checkAutoNodeLabelrefresh
  checkJournalNodeSync
  checkCheckpoint
  checkCleanAcls
  checkSsdMonitor
  checkOperationCollector
  checkMapReduceDistributedCache
  #checkBlockplacementLog
  checkAutoDiskBalancer
}
```

- Step 3** Save the file, log in to Manager, and choose **Cluster > Services > HDFS > Instance**. Select all NameNode instances, click **More**, and choose **Restart Instance**.

----End

## 12.35 ALM-14012 JournalNode Is Out of Synchronization Is Generated in the Cluster

### Symptom

Alarm "ALM-14012 JournalNode Is Out of Synchronization" is generated in the MRS cluster.

### Cause Analysis

1. Log in to the node for which the alarm is generated and search for the **startDetail.log** file of the JournalNode instance in the **/var/log/Bigdata/hdfs/nn** directory. The log shows that the JournalNode instance has stopped.
2. Check the latest **edits** log file in the **/srv/BigData/journalnode/hacluster/current** directory on the node for which the alarm is generated and other JournalNode nodes. It is found that data on the faulty node is not synchronized with other nodes.

### Procedure

- Step 1** Log in to FusionInsight Manager and choose **Cluster > Services > HDFS > Instance**. On the displayed page, select the JournalNode instance of the node for which the alarm is generated, click **More**, and select **Stop Instance**.
- Step 2** Log in to the node for which the alarm is generated, and move all files in the **/srv/BigData/journalnode/hacluster/current** directory to another new directory (for example, **/opt/test**) to clear the directory.
- Step 3** Log in to FusionInsight Manager and choose **Cluster > Services > HDFS > Instance**. On the displayed page, select the stopped JournalNode instance and click **Start Instance**.
- Step 4** Wait for a while and check whether the alarm is cleared.

----End

## 12.36 Failed to Decommission a DataNode Due to HDFS Block Loss

### Symptom

A DataNode fails to be decommissioned.

### Cause Analysis

1. Check the decommissioning log. It shows that there are 1564 blocks but one block cannot be backed up.

```
[2021-02-25 09:52:46]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:52:50]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:52:55]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:52:59]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:53:03]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:53:07]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:53:11]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:53:16]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:53:20]NameNode#192.168.1.51, current underReplicatedBlocks: 1, total blocks: 1564.
[2021-02-25 09:53:20]Node 192.168.1.44 decommission failed, no data was decommissioned in 30 minutes.
[2021-02-25 09:53:20]192.168.1.44#DataNode cannot be decommissioned.
```

2. Log in to the master node of the cluster, go to the HDFS client, run the **hdfs fsck /** command to check the damaged block, and record the file path.

Example: **/tmp/hive-scratch/omm/\_tez\_session\_dir/xxx-resources/xxx.jar**

The HDFS status is **CORRUPT**.

```
Erasure Coded Block Groups:
Total size: 0 B
Total files: 0
Total block groups (validated): 0
Minimally erasure-coded block groups: 0
Over-erasure-coded block groups: 0
Under-erasure-coded block groups: 0
Unsatisfactory placement block groups: 0
Average block group size: 0.0
Missing block groups: 0
Corrupt block groups: 0
Missing internal blocks: 0
FSCK ended at Thu Feb 25 17:36:56 CST 2021 in 2584 milliseconds

The filesystem under path '/' is CORRUPT
```

## Procedure

**Step 1** Check whether the damaged block can be deleted.

- If yes, go to **Step 2**.
- If no, contact technical support.

**Step 2** Run the following commands to log in to the HDFS client:

```
cd HDFS client installation directory
```

```
source bigdata_env
```

```
kinit Service user
```

**Step 3** Run the following command to delete the damaged block:

```
hdfs dfs -rm -skipTrash /tmp/hive-scratch/omm/_tez_session_dir/xxx-resources/xxx.jar
```

**Step 4** Run the following command to check whether the HDFS status is restored to **HEALTHY**:

```
hdfs fsck /
```

```
Erasure Coded Block Groups:
Total size:      0 B
Total files:     0
Total block groups (validated):      0
Minimally erasure-coded block groups: 0
Over-erasure-coded block groups:     0
Under-erasure-coded block groups:     0
Unsatisfactory placement block groups: 0
Average block group size:             0.0
Missing block groups:                 0
Corrupt block groups:                 0
Missing internal blocks:              0
FSCK ended at Thu Feb 25 17:38:38 CST 2021 in 2555 milliseconds

The filesystem under path '/' is HEALTHY
```

**Step 5** Decommission the DataNode again.

----End

## 12.37 An Error Is Reported When DistCP Is Used to Copy an Empty Folder

### Symptom

When a user runs the following **distcp** commands on the MRS client, empty folders cannot be copied from HDFS to OBS:

```
hadoop distcp -Dfs.obs.endpoint=xxx
```

```
-Dfs.obs.access.key=xxx -Dfs.obs.secret.key=xxx -update hdfs://hacluster/blee
obs://xxx/aaa
```

### Cause Analysis

If the source directory (for example, **blee**) is empty and the destination directory (for example, **aaa**) does not exist, the system automatically creates the **aaa** directory but does not create the **blee** directory under **aaa**.

If the **blee** directory at the source is not empty and the **aaa** directory at the destination does not exist, the system automatically creates the **aaa** directory and creates the **blee** directory under **aaa** to move files.

### Procedure

- It is recommended that the source directory (for example, **blee**) is not empty.
- If the source directory is empty, manually create a destination directory (**aaa** for example) before moving files.

# 13 Using Hive

## 13.1 Common Hive Logs

### Audit log

An audit log records at what time a user sends a request to HiveServer and MetaStore from which IP address with what statement.

The following HiveServer audit log shows that at 14:51:22 on February 1, 2016, **user\_chen** sent a **show tables** request to HiveServer from the 192.168.1.18 IP address.

```
2016-02-01 14:51:22,335 | INFO | HiveServer2-Handler-Pool: Thread-37815 | UserN  
ame=user_chen | ip=192.168.1.18 | Time=2016/02/01 14:51:22 | Opertai  
on=ExecuteStatement | stmt={show tables} | Resource= | Result= Detail=  
| org.apache.hive.service.cli.thrift.ThriftCLIService.logAuditEvent(ThriftCLISer  
vice.java:350)
```

The following MetaStore audit log shows that user **hive** sent a **shutdown** request to MetaStore from IP address 192.168.1.18 at 11:31:15 on January 29, 2016.

```
2016-01-29 11:31:15,451 | INFO | pool-6-thread-70648 | ugi=hive/hadoop.hadoop.c  
om@HADOOP.COM | ip=192.168.1.18 | cmd=Shutting down the object store...  
| org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.logAuditEvent(HiveM  
etaStore.java:375)
```

The audit log must be checked to solve the following problems:

1. There is no response after a client sends a request. The audit log can be used to check whether the task suspends on the client or server. If the audit log does not contain related information, the task is terminated on the client. If the audit log contains related information, locate the fault according to the run log.
2. The audit log can be used to check the number of requests in a specified period of time.

### HiveServer Run Log

HiveServer receives requests from a client (SQL statement), compile and execute the statement (submitted to Yarn or local MapReduce), and interact with

MetaStore to obtain metadata information. The HiveServer run log records a complete SQL execution process.

Generally, if a SQL statement fails to be executed, view the HiveServer run log first.

## MetaStore Run Log

Typically, if the HiveServer run log contains a MetaException or MetaStore connection failure, check the MetaStore run log.

## GC Log

Both HiveServer and MetaStore have GC logs. If GC-related problems occur, view the GC logs to quickly locate the cause. For example, if HiveServer or MetaStore frequently restarts, check its GC log.

# 13.2 Failed to Start Hive

The most common cause of the Hive startup failure is that the MetaStore instance cannot connect to DBService. The error information is printed in the MetaStore log.

## Possible Causes

1. DBService does not properly initialize the Hive metabase hivemeta.
2. The floating IP address of DBService is incorrect. As a result, the IP address of the MetaStore node fails to connect to or build mutual trust with the floating IP address, causing the MetaStore startup failure.

## Procedure 1

**Step 1** Run the following commands:

```
source /opt/Bigdata/MRS_XXX/install/dbservice/.dbservice_profile
```

```
gsql -h DBService floating IP address -p 20051 -d hivemeta -U hive -W Hive user password
```

**Step 2** If the interaction interface cannot be properly displayed, database initialization fails. If the following error information is displayed, the hivemeta configuration may be lost in the configuration file of the node where DBService is located.

```
org.postgresql.util.PSQLException: FATAL: no pg_hba.conf entry for host "192.168.0.146", database "HIVEMETA"
```

**Step 3** Edit `/srv/BigData/dbdata_service/data/pg_hba.conf` by adding **host hivemeta hive 0.0.0.0/0 sha256** to the file.

**Step 4** Run the **source /opt/Bigdata/MRS\_XXX/install/dbservice/.dbservice\_profile** command to configure environment variables.

**Step 5** Run **gs\_ctl -D \$GAUSSDATA reload #** to make new configurations take effect.

----End

## Procedure 2

The floating IP address of DBService must be an IP address that is not used in the same network segment and cannot be pinged before configuration. Change the floating IP address of DBService.

# 13.3 Error Message "Cannot modify xxx at runtime" Is Displayed When the set Command Is Executed in a Security Cluster

## Symptom

The following error is reported when running the `set` command:

```
O: jdbc:hive2://192.168.1.18:21066/> set mapred.job.queue.name=QueueA;
Error: Error while processing statement: Cannot modify mapred.job.queue.name at list of params that are allowed to be modified at runtime (state=42000,code=1)
```

## Procedure

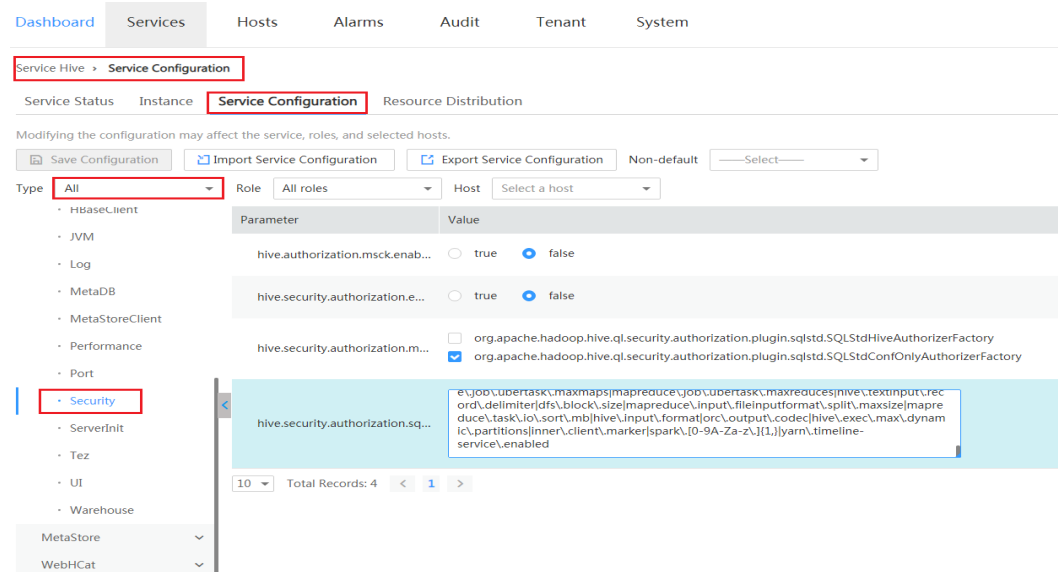
### Solution 1:

**Step 1** Log in to Manager and modify Hive parameters.

- MRS Manager: Log in to MRS Manager and choose **Services > Hive > Service Configuration**. Set **Type** to **All** and choose **HiveServer > Security**.
- FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Hive**. On the page that is displayed, choose **Configurations > All Configurations > HiveServer > Security**.

**Step 2** Add the command parameters to be executed to the `hive.security.authorization.sqlstd.confwhitelist.append` configuration item.

**Step 3** Click **Save** and restart **HiveServer**.



----End

**Solution 2:**

**Step 1** Log in to Manager and modify Hive parameters.

- MRS Manager: Log in to MRS Manager and choose **Services > Hive > Service Configuration**. Set **Type** to **All** and choose **HiveServer > Security**.
- FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Hive**. On the page that is displayed, choose **Configurations > All Configurations > HiveServer > Security**.

**Step 2** Locate **hive.security.whitelist.switch** and select **OFF**. Click **Save** and restart HiveServer.

----End

## 13.4 Specifying a Queue When Submitting a Hive Task

### Symptom

How do I specify a queue when Hive submits a job?

### Procedure

**Step 1** Before submitting the job, set the job queue, for example, submitting the job to QueueA.

```
set mapred.job.queue.name=QueueA;
```

```
select count(*) from rc;
```

 **NOTE**

The queue name is case sensitive. For example, in this example, **queueA** and **Queuea** are invalid. In addition, the queue must be a leaf queue, and jobs cannot be submitted to a non-leaf queue.

**Step 2** After job submission, go to the Yarn page to check the job. The job has been submitted to QueueA.

<b>User:</b>	admin
<b>Name:</b>	select count(*) from rc(Stage-1)
<b>Application Type:</b>	MAPREDUCE
<b>Application Tags:</b>	
<b>YarnApplicationState:</b>	FINISHED
<b>Queue:</b>	QueueA
<b>FinalStatus Reported by AM:</b>	SUCCEEDED
<b>Started:</b>	Thu Mar 03 09:01:58 +0800 2016
<b>Elapsed:</b>	1mins, 0sec
<b>Tracking URL:</b>	History
<b>Log Aggregation Status</b>	Status
<b>Diagnostics:</b>	

----End



## 13.5 Setting the Map/Reduce Memory on the Client

### Symptom

How do I set Map and Reduce memory on the client?

### Procedure

Before SQL statement execution, run the **set** command to set parameters of clients related to Map/Reduce.

The following parameters are related to Map and Reduce memory:

```
set mapreduce.map.memory.mb=4096; //Memory required by each Map task
set mapreduce.map.java.opts=-Xmx3276M; //Maximum memory used by the JVM of each Map task
set mapreduce.reduce.memory.mb=4096; //Memory required by each Reduce task
set mapreduce.reduce.java.opts=-Xmx3276M; //Maximum memory used by the JVM of each Reduce task
set mapred.child.java.opts=-Xms1024M -Xmx3584M; // This parameter is a global parameter, which is used
to set Map and Reduce in a unified manner.
```

#### NOTE

Parameter settings take effect for the current session only.

## 13.6 Specifying the Output File Compression Format When Importing a Hive Table

### Symptom

The user does not know how to specify an output file compression format when importing a Hive table.

### Procedure

Hive supports the following compression formats:

```
org.apache.hadoop.io.compress.BZip2Codec
org.apache.hadoop.io.compress.Lz4Codec
org.apache.hadoop.io.compress.DeflateCodec
org.apache.hadoop.io.compress.SnappyCodec
org.apache.hadoop.io.compress.GzipCodec
```

- If global settings are required, that is, all tables need to be compressed, you can perform the following global settings for Hive service configuration parameters on the Manager page:
  - Set **hive.exec.compress.output** to **true**.
  - Set **mapreduce.output.fileoutputformat.compress.codec** to **org.apache.hadoop.io.compress.BZip2Codec**.

#### NOTE

The following parameters take effect only when **hive.exec.compress.output** is set to **true**.

- If it needs to be set at the session level, configure the parameters as follows before command execution:

```
set hive.exec.compress.output=true;
set
mapreduce.output.fileoutputformat.compress.codec=org.apache.hadoop.io.compress.SnappyCodec;
```

## 13.7 Description of the Hive Table Is Too Long to Be Completely Displayed

### Symptom

How do I make sure that the description is completely displayed when the desc table is too long?

### Procedure

**Step 1** When starting Beeline of Hive, set **maxWidth** to **20000**.

```
[root@192-168-1-18 logs]# beeline --maxWidth=20000
scan complete in 3ms
Connecting to
...
Beeline version 1.1.0 by Apache Hive
```

**Step 2** (Optional) Run the **beeline -help** command to view the client display settings.

```
-u <database url>          the JDBC URL to connect to
-n <username>              the username to connect as
-p <password>              the password to connect as
-d <driver class>          the driver class to use
-i <init file>             script file for initialization
-e <query>                 query that should be executed
-f <exec file>             script file that should be executed
--hiveconf property=value Use value for given property
--color=[true/false]      control whether color is used for display
--showHeader=[true/false] show column names in query results
--headerInterval=ROWS;    the interval between which headers are displayed
--fastConnect=[true/false] skip building table/column list for tab-completion
--autoCommit=[true/false] enable/disable automatic transaction commit
--verbose=[true/false]    show verbose error messages and debug info
--showWarnings=[true/false] display connection warnings
--showNestedErrs=[true/false] display nested errors
--numberFormat=[pattern]  format numbers using DecimalFormat pattern
--force=[true/false]      continue running script even after errors
--maxWidth=MAXWIDTH        the maximum width of the terminal
--maxColumnWidth=MAXCOLWIDTH the maximum width to use when displaying columns
--silent=[true/false]     be more silent
--autosave=[true/false]   automatically save preferences
--outputformat=[table/vertical/csv2/tsv2/dsv/csv/tsv] format mode for result display
Note that csv, and tsv are deprecated - use csv2, tsv2 instead
--truncateTable=[true/false] truncate table column when it exceeds length
--delimiterForDSV=DELIMITER specify the delimiter for delimiter-separated values output format
(default: |)
--isolation=LEVEL         set the transaction isolation level
--nullemptystring=[true/false] set to true to get historic behavior of printing null as empty string
--socketTimeOut=n        socket connection timeout interval, in second. The default value is 300.
```

----End

## 13.8 NULL Is Displayed When Data Is Inserted After the Partition Column Is Added to a Hive Table

### Symptom

1. Run the following command to create a table:

```
create table test_table(  
col1 string,  
col2 string  
)  
PARTITIONED BY(p1 string)  
STORED AS orc tblproperties('orc.compress'='SNAPPY');
```
2. Modify the table structure, add partitions, and insert data.

```
alter table test_table add partition(p1='a');  
insert into test_table partition(p1='a') select col1,col2 from temp_table;
```
3. Modify the table structure, add columns, and insert data.

```
alter table test_table add columns(col3 string);  
insert into test_table partition(p1='a') select col1,col2,col3 from temp_table;
```
4. Query data in the **test\_table** table. In the returned result, the values in the **col3** column are all NULL.

```
select * from test_table where p1='a'
```
5. Add a table partition and insert data.

```
alter table test_table add partition(p1='b');  
insert into test_table partition(p1='b') select col1,col2,col3 from temp_table;
```
6. Query data in the **test\_table** table. In the returned result, the value of **col3** is not all NULL.

```
select * from test_table where p1='b'
```

### Cause Analysis

**RESTRICT** is the default option for altering a table. In the RESTRICT mode, only the metadata is changed, while the table's partition structure created before the altering operation remains unchanged. However, new partitions created after the altering operation are changed. Therefore, when values of the old partitions are queried, they are all NULL.

### Procedure

Add the **cascade** keyword when adding columns, for example:

```
alter table test_table add columns(col3 string) cascade;
```

## 13.9 New User Created in the Cluster Does Not Have the Permission to Query Hive Data

### Symptom

When a user is created, an error message is displayed indicating that the user does not have permissions to query data.

```
Error: Error while compiling statement: FAILED: HiveAccessControlException Permission denied: Principal [name=hive, type=USER] does not have following privileges for operation QUERY [[SELECT] on Object [type=TABLE_OR_VIEW, name=default.t1]] (state=42000,code=40000)
```

## Cause Analysis

The newly created user does not have the permission to operate the Hive component.

## Solution

MRS Manager:

**Step 1** Log in to MRS Manager and choose **System > Manage Role > Create Role**.

**Step 2** Enter a role name.

**Step 3** In the **Permission** area, select **Hive**. The Hive administrator permission and the read and write permission for Hive tables are displayed.

The screenshot shows the 'Create Role' configuration page. The 'Role Name' is 'hive\_user'. Under the 'Permission' section, the service is set to 'Hive'. A table of permissions is displayed, with 'Hive Read Write Privileges' selected and highlighted by a red box. Below it, 'Hive Admin Privilege' is not selected. The bottom of the table shows a pagination control with '10' records per page, 'Total Records: 2', and page '1' of 1.

**Step 4** Select **Hive Read Write Privileges**. All databases in the Hive column are displayed.

**Step 5** Select the permissions required by the role and click **OK**.

**Step 6** On MRS Manager, choose **System > Manage User**.

**Step 7** Locate the row that contains the created user, and click **Modify** in the **Operation** column.

**Step 8** Click **Select and Join User Group**. To use the Hive service, you must add a Hive group.

**Step 9** Click **Select and Add Role** and select the role created in [Step 5](#).

**Step 10** Click **OK**.

----End

FusionInsight Manager:

- Step 1** Log in to FusionInsight Manager. Choose **System > Permission > Role**.
  - Step 2** Click **Create Role**, and set **Role name** and **Description**.
  - Step 3** Set **Configure Resource Permission** for the role and select **Hive Read and Write Permission** for the Hive table. All databases in the Hive column are displayed.
  - Step 4** Select the permissions required by the role and click **OK**.
  - Step 5** On FusionInsight Manager, choose **System > Permission > User**.
  - Step 6** Locate the row that contains the created user, and click **Modify** in the **Operation** column.
  - Step 7** Click **Add** on the right of **User Group**. To use the Hive service, you must add a Hive group.
  - Step 8** Click **Add** on the right of **Role** and select the role created in [4](#).
  - Step 9** Click **OK**.
- End

## 13.10 An Error Is Reported When SQL Is Executed to Submit a Task to a Specified Queue

### Symptom

The following error is reported when executing SQL to submit a task to Yarn:

```
Failed to submit application_1475400939788_0033 to YARN :  
org.apache.hadoop.security.AccessControlException: User newtest cannot submit applications to queue  
root.QueueA
```

### Cause Analysis

The current login user does not have the permission to submit the YARN queue.

### Solution

Grant the submission permission of the specified Yarn queue to the user. On Manager, choose **System > Permission > User** and bind a role with the queue submission permission to the user.

## 13.11 An Error Is Reported When the "load data inpath" Command Is Executed

### Symptom

The following errors are reported when the **load data inpath** command is executed:

- Error 1:

```
HiveAccessControlException Permission denied. Principal [name=user1, type=USER] does not have following privileges on Object [type=DFS_URI, name=hdfs://hacluster/tmp/input/mapdata] for operation LOAD : [OBJECT OWNERSHIP]
```

- Error 2:

```
HiveAccessControlException Permission denied. Principal [name=user1, type=USER] does not have following privileges on Object [type=DFS_URI, name=hdfs://hacluster/tmp/input/mapdata] for operation LOAD : [INSERT, DELETE]
```

- Error 3:

```
SemanticException [Error 10028]: Line 1:17 Path is not legal "file:///tmp/input/mapdata": Move from: file:/tmp/input/mapdata to: hdfs://hacluster/user/hive/warehouse/tmp1 is not valid. Please check that values for params "default.fs.name" and "hive.metastore.warehouse.dir" do not conflict.
```

## Cause Analysis

The current login user does not have the permission to operate the directory or the file directory format is incorrect.

## Solution

Hive has the following requirements on the **load data inpath** command:

- The file owner must be the user who executes the command.
- The current user must have read and write permissions for the file.
- The current user must have permissions to execute the directory of the file.
- The current user must have the write permission on the directory of the table, because the load operation moves the file to the directory.
- The file format must be the same as the storage format specified by the table. For example, if **stored as rcfile** is specified during table creation but the file format is TXT, it is unsatisfied.
- The file must be stored in HDFS. Files in the local file system cannot be specified using the **file://** form.
- The file name cannot start with an underscore (`_`) or period (`.`). A file whose name starts with an underscore (`_`) or period (`.`) will be ignored.

The following shows permissions required when user **test\_hive** loads data.

```
[root@192-168-1-18 duan]# hdfs dfs -ls /tmp/input2
16/03/21 14:45:07 INFO hdfs.PeerCache: SocketCache disabled.
Found 1 items
-rw-r--r--  3 test_hive hive      6 2016-03-21 14:44 /tmp/input2/input.txt
```

## 13.12 An Error Is Reported When the "load data local inpath" Command Is Executed

### Symptom

The following errors are reported when the **load data local inpath** command is executed:

- Error 1:

```
HiveAccessControlException Permission denied. Principal [name=user1, type=USER] does not have following privileges on Object [type=LOCAL_URI, name=file:/tmp/input/mapdata] for operation LOAD : [SELECT, INSERT, DELETE]
```

- Error 2:

```
HiveAccessControlException Permission denied. Principal [name=user1, type=USER] does not have following privileges on Object [type=LOCAL_URI, name=file:/tmp/input/mapdata] for operation LOAD : [OBJECT OWNERSHIP]
```

- Error 3:  
SemanticException Line 1:23 Invalid path "/tmp/input/mapdata": No files matching path file:/tmp/input/mapdata

## Cause Analysis

The current user does not have the permission to operate the directory or the directory does not exist on the node where HiveServer is located.

## Solution

### NOTE

Generally, you are not advised to use local files to load data to Hive tables. You are advised to store local files in HDFS and then load data from the cluster.

Hive has the following requirements on the **load data local inpath** command:

- The file must be stored on the HiveServer node, because all commands are sent to the active HiveServer for execution.
- User **omm** must have the read permission for the file and read and execution permissions for the directory where the file is located, because the HiveServer process is started by user **omm** in the OS.
- The file owner must be the user who executes the command.
- The current user must have read and write permissions for the file.
- The file format must be the same as the storage format specified by the table. For example, if **stored as rcfile** is specified during table creation but the file format is TXT, it is unsatisfied.
- The file name cannot start with an underscore (`_`) or period (`.`). A file whose name starts with an underscore (`_`) or period (`.`) will be ignored.

## 13.13 An Error Is Reported When the create external table Command Is Executed

### Symptom

The following error is reported when the **create external table *xx(xx int) stored as textfile location '/tmp/aaa/aaa'*** command is executed.

```
Permission denied. Principal [name=fantasy, type=USER] does not have following privileges on Object [type=DFS_URI, name=/tmp/aaa/aaa] for operation CREATETABLE : [SELECT, INSERT, DELETE, OBJECT OWNERSHIP] (state=42000,code=40000)
```

### Cause Analysis

The current login user does not have the read and write permissions for the directory or its parent directory. When an external table is created, whether the current user is checked for its read and write permissions for the specified directory and its subdirectories and subfiles. If the specified directory does not exist, permissions for the parent directory are checked, and so on. If the check

results show that the user has no permissions on any directory, "insufficient permission" is reported instead of "The specified directory does not exist".

## Solution

Check whether the current user has read and write permissions for the `/tmp/aaa/aaa` path. If the path does not exist, check whether the user has read and write permissions for its parent directory.

## 13.14 An Error Is Reported When the `dfs -put` Command Is Executed on the Beeline Client

### Symptom

Run the following command:

```
dfs -put /opt/kv1.txt /tmp/kv1.txt
```

The following error is reported:

```
Permission denied. Principal [name=admin, type=USER] does not have following privileges onObject[type=COMMAND_PARAMS,name=[-put, /opt/kv1.txt, /tmp/kv1.txt]] for operation DFS : [ADMIN PRIVILEGE] (state=,code=1)
```

### Cause Analysis

The current login user does not have the permissions to run the command.

### Solution

If the current user has the **admin** role, run the **set role admin** command to switch to the **admin** role. If the user does not have the admin role, bind the user with the permissions of the corresponding role on the Manager page.

## 13.15 Insufficient Permissions to Execute the `set role admin` Command

### Symptom

When a user runs the following command:

```
set role admin
```

The following error is reported:

```
0: jdbc:hive2://192.168.42.26:21066/> set role admin;  
Error: Error while processing statement: FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask. dmp_B doesn't belong to role admin (state=08S01,code=1)
```

### Cause Analysis

The current user does not have the permissions of the **admin** role of Hive.



## Solution

- Step 1** Log in to Manager.
- For versions earlier than MRS 3.x, go to [Step 7](#).
  - For MRS 3.x or later, choose **Cluster > Services > Hive**. In the upper right corner of the **Dashboard** page, click **More** and check whether **Enable Ranger** is unavailable.
    - If yes, go to [Step 2](#).
    - If no, go to [Step 7](#).
- Step 2** Choose **Cluster > Services > Ranger** and click **RangerAdmin** in the **Basic Information** area. The Ranger web UI is displayed.
- Step 3** Click the username in the upper right corner, select **Log Out** to log out of the system, and log in to the system as user **rangeradmin**.
- Step 4** On the homepage, click **Settings** and choose **Roles**.
- Step 5** Click the role with **Role Name** set to **admin**. In the **Users** area, click **Select User** and select a username.
- Step 6** Click **Add Users**, select **Is Role Admin** in the row where the username is located, and click **Save**.
- Step 7** Choose **System > Permission > Role** and add a role with the Hive administrator permission.
- Step 8** On FusionInsight Manager, choose **System > Permission > User**.
- Step 9** In the **Operation** column of the user, click **Modify**.
- Step 10** Bind a role that has the Hive administrator permissions to the user and click **OK**.
- End

## 13.16 An Error Is Reported When a UDF Is Created on the Beeline Client

### Symptom

Run the following command:

```
create function fn_test3 as 'test.MyUDF' using jar 'hdfs:///tmp/udf2/MyUDF.jar'
```

The following error is reported:

```
Error: Error while compiling statement: FAILED: HiveAccessControlException Permission denied: Principal [name=admin, type=USER] does not have following privileges for operation CREATEFUNCTION [[ADMIN PRIVILEGE] on Object [type=DATABASE, name=default], [ADMIN PRIVILEGE] on Object [type=FUNCTION, name=default.fn_test3]] (state=42000,code=40000)
```

### Cause Analysis

To create a permanent function in Hive, role **admin** is required.

## Solution

Run the **set role admin** command before running the statement.

## 13.17 Hive Is Faulty

### Possible Causes

- The DBService service is not available.
- The HDFS service is not available.
- The ZooKeeper service is not available.
- The LDAP/KrbServer service is not available.
- The MetaStore instance is not available.

### Solution

- If the DBService is unavailable, see [ALM-27001 DBService Is Unavailable](#).
- If the HDFS service is unavailable, see [ALM-14000 HDFS Service Unavailable](#).
- If the ZooKeeper service is unavailable, see [ALM-13000 ZooKeeper Service Unavailable](#).
- If the LDAP/KrbServer service is unavailable, see [ALM-25000 LdapServer Service Unavailable/ALM-25500 KrbServer Service Unavailable](#).
- If the Metastore instance is unavailable, see [ALM-16004 Hive Service Unavailable](#).

## 13.18 Difference Between Hive Service Health Status and Hive Instance Health Status

### Question

What is the difference between Hive service health status and Hive instance health status?

### Solution

The Hive service health status is displayed on Manager and has four values: **Good**, **Bad**, **Partially Healthy**, and **Unknown**. It depends not only on Hive service availability but also the service status of other related components. Simple SQL is used to check Hive service availability.

Hive instances consist of HiveServer and MetaStore. Their health status is determined by communications between instances and JMX and can be **Good** (normal communications), **Concerning** (abnormal communications), or **Unknown** (no communications).

## 13.19 "authentication failed" Is Displayed During an Attempt to Connect to the Shell Client

### Symptom

In clusters in security mode, the **beeline** command fails to be executed on the Shell client when the HiveServer service is normal, and the system prompts "authentication failed". The following information is displayed.

```
Debug is true storeKey false useTicketCache true useKeyTab false doNotPrompt false ticketCache is null
isInitiator true KeyTab is null refreshKrb5Config is false principal is null tryFirstPass is false useFirstPass is
false storePass is false clearPass is false
Acquire TGT from Cache
Credentials are no longer valid
Principal is null
null credentials from Ticket Cache
[Krb5LoginModule] authentication failed
No password provided
```

### Cause Analysis

- The client user does not perform security authentication.
- Kerberos authentication expired.

### Solution

**Step 1** Log in to the node where the Hive client is installed.

**Step 2** Run the **source** *Cluster client installation directory*/**bigdata\_env** command.

Run the **klist** command to check whether there is a valid ticket in the local end. The following information shows that the ticket became valid at 14:11:42 on December 24, 2016, and expired at 14:11:40 on December 25, 2016. In the period of time, the ticket was available.

```
klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: xxx@HADOOP.COM
Valid starting Expires Service principal
12/24/16 14:11:42 12/25/16 14:11:40 krbtgt/HADOOP.COM@HADOOP.COM
```

**Step 3** Run the **kinit** *username* command for authentication and log in to the client again.

----End

## 13.20 Failed to Access ZooKeeper from the Client

### Symptom

In clusters in security mode, when the HiveServer service is normal and SQL is executed by using the JDBC interface to connect to HiveServer, "The ZooKeeper client is AuthFailed" is reported.

```
14/05/19 10:52:00 WARN utils.HAClientUtilDummyWatcher: The ZooKeeper client is AuthFailed
14/05/19 10:52:00 INFO utils.HiveHAClientUtil: Exception thrown while reading data from znode.The
```

```
possible reason may be connectionless. This is recoverable. Retrying..
14/05/19 10:52:16 WARN utils.HAClientUtilDummyWatcher: The ZooKeeper client is AuthFailed
14/05/19 10:52:32 WARN utils.HAClientUtilDummyWatcher: The ZooKeeper client is AuthFailed
14/05/19 10:52:32 ERROR st.BasicTestCase: Exception: Could not establish connection to active hiveserver
java.sql.SQLException: Could not establish connection to active hiveserver
```

Or an error is reported stating "Unable to read HiveServer2 configs from ZooKeeper":

```
Exception in thread "main" java.sql.SQLException: org.apache.hive.jdbc.ZooKeeperHiveClientException:
Unable to read HiveServer2 configs from ZooKeeper
at org.apache.hive.jdbc.HiveConnection.<init>(HiveConnection.java:144)
at org.apache.hive.jdbc.HiveDriver.connect(HiveDriver.java:105)
at java.sql.DriverManager.getConnection(DriverManager.java:664)
at java.sql.DriverManager.getConnection(DriverManager.java:247)
at JDBCExample.main(JDBCExample.java:82)
Caused by: org.apache.hive.jdbc.ZooKeeperHiveClientException: Unable to read HiveServer2 configs from
ZooKeeper
at
org.apache.hive.jdbc.ZooKeeperHiveClientHelper.configureConnParams(ZooKeeperHiveClientHelper.java:100)
at org.apache.hive.jdbc.Utils.configureConnParams(Utils.java:509)
at org.apache.hive.jdbc.Utils.parseURL(Utils.java:429)
at org.apache.hive.jdbc.HiveConnection.<init>(HiveConnection.java:142)
... 4 more
Caused by: org.apache.zookeeper.KeeperException$ConnectionLossException: KeeperErrorCode =
ConnectionLoss for /hiveserver2
at org.apache.zookeeper.KeeperException.create(KeeperException.java:99)
at org.apache.zookeeper.KeeperException.create(KeeperException.java:51)
at org.apache.zookeeper.ZooKeeper.getChildren(ZooKeeper.java:2374)
at org.apache.curator.framework.imps.GetChildrenBuilderImpl$3.call(GetChildrenBuilderImpl.java:214)
at org.apache.curator.framework.imps.GetChildrenBuilderImpl$3.call(GetChildrenBuilderImpl.java:203)
at org.apache.curator.RetryLo, op.callWithRetry(RetryLoop.java:107)
at
org.apache.curator.framework.imps.GetChildrenBuilderImpl.pathInForeground(GetChildrenBuilderImpl.java:2
00)
at org.apache.curator.framework.imps.GetChildrenBuilderImpl.forPath(GetChildrenBuilderImpl.java:191)
at org.apache.curator.framework.imps.GetChildrenBuilderImpl.forPath(GetChildrenBuilderImpl.java:38)
```

## Cause Analysis

- When the client connects to HiveServer, the HiveServer address is automatically obtained from ZooKeeper. If ZooKeeper connection authentication is abnormal, the HiveServer address cannot be obtained from ZooKeeper correctly.
- During ZooKeeper connection authentication, **krb5.conf**, **principal**, **keytab**, and related information must be loaded to the client. Authentication failure causes are as follows:
  - The **user.keytab** path is incorrectly entered.
  - **user.principal** is incorrectly entered.
  - The cluster has switched the domain name. However, the old principal is used when the client combines the URL.
  - The client cannot pass Kerberos authentication due to firewall settings. Ports 21730 (TCP), 21731 (TCP/UDP), and 21732 (TCP/UDP) need to be opened for Kerberos.

## Solution

- Step 1** Ensure that the user can properly access the **user.keytab** file in related paths on the client node.

**Step 2** Ensure that the user's **user.principal** corresponds to the specified **keytab** file.

Run the **klist -kt keytabpath/user.keytab** command to check the file.

**Step 3** If the cluster has switched the domain name, the **principal** field used in the URL must be the new domain name.

For example, the default value is **hive/hadoop.hadoop.com@HADOOP.COM**. If the cluster has switched the domain name, the field must be changed accordingly. For example, if the domain name is **abc.com**, enter **hive/hadoop.abc.com@ABC.COM**.

**Step 4** Ensure that authentication is normal and HiveServer can be connected.

Run the following commands on the client:

```
source Client installation directory/bigdata_env
```

```
kinit username
```

Run the **beeline** command on the client to ensure normal running.

----End

## 13.21 "Invalid function" Is Displayed When a UDF Is Used

### Symptom

When a UDF is created on the Hive client using Spark, "Error 10011" indicating "invalid function" is reported:

```
Error: Error while compiling statement: FAILED: SemanticException [Error 10011]: Line 1:7 Invalid function 'test_udf' (state=42000,code=10011)
```

The preceding problem occurs when multiple HiveServers use a UDF. For example, if metadata is not synchronized in time when the UDF created on HiveServer2 is used on HiveServer1, the preceding error is reported when clients on HiveServer1 are connected.

### Cause Analysis

Metadata shared by multiple HiveServers or Hive and Spark is not synchronized, causing memory data inconsistency between different HiveServer instances and invalid UDF.

### Solution

Synchronize new UDF information to HiveServer and reload the function.

## 13.22 Hive Service Status Is Unknown

### Symptom

The Hive service status is unknown.

## Cause Analysis

The Hive service stops.

## Solution

Restart the Hive service.

## 13.23 Health Status of a HiveServer or MetaStore Instance Is unknown

### Symptom

The health status of a HiveServer or MetaStore instance is unknown.

### Cause Analysis

The HiveServer or MetaStore instance is stopped.

### Solution

Restart the HiveServer or MetaStore instance.

## 13.24 Health Status of a HiveServer or MetaStore Instance Is Concerning

### Symptom

The health status of the HiveServer or MetaStore instance is **Concerning**.

### Cause Analysis

The HiveServer or MetaStore instance cannot be started properly. For example, when modifying the MetaStore/HiveServer GC parameter, you can view the startup log of the corresponding process, for example, the **hiveserver.out(hadoop-omm-jar-192-168-1-18.out)** file. The following exception occurs:

```
Error: Could not find or load main class Xmx2048M
```

The preceding information indicates that **Xmx2048M** is used as the startup parameter of the Java process instead of the JVM during the startup of the JVM. As shown in the following information, the hyphen (-) is deleted mistakenly.

```
METASTORE_GC_OPTS=Xms1024M Xmx2048M -DignoreReplayReqDetect  
-XX\;CMSFullGCsBeforeCompaction\=1 -XX\;+UseConcMarkSweepGC  
-XX\;+CMSParallelRemarkEnabled -XX\;+UseCMSCompactAtFullCollection  
-XX\;+ExplicitGCInvokesConcurrent -server -XX\;MetaspaceSize\=128M  
-XX\;MaxMetaspaceSize\=256M
```

## Solution

Check the latest changes to detect incorrect settings.

```
METASTORE_GC_OPTS=Xms1024M -Xmx2048M -DignoreReplayReqDetect  
-XX\:CMSFullGCsBeforeCompaction\=1 -XX\:+UseConcMarkSweepGC  
-XX\:+CMSParallelRemarkEnabled -XX\:+UseCMSCompactAtFullCollection  
-XX\:+ExplicitGCInvokesConcurrent -server -XX\:MetaspaceSize\=128M  
-XX\:MaxMetaspaceSize\=256M
```

## 13.25 Garbled Characters Returned Upon a Query If Text Files Are Compressed Using ARC4

### Symptom

If a Hive query result table is compressed and stored using the ARC4 algorithm, garbled characters are returned after the select \* query is conducted in the result table.

### Cause Analysis

The default Hive compression format is not ARC4 or output compression is disabled.

### Solution

**Step 1** If garbled characters are returned after the SETECT query, set the following in Beeline:

```
set  
mapreduce.output.fileoutputformat.compress.codec=org.apache.hadoop.io.enc  
ryption.arc4.ARC4BlockCodec;  
  
set hive.exec.compress.output=true;
```

**Step 2** Import the table to a new table using block decompression.

```
insert overwrite table tbl_result select * from tbl_source;
```

**Step 3** Perform the query again.

```
select * from tbl_result;  
  
----End
```

## 13.26 Hive Task Failed to Run on the Client but Successful on Yarn

### Symptom

When a Hive task fails to be executed, an error message similar to the following is displayed on the client:

Error:Invalid OperationHandler:OperationHander [opType=EXECUTE\_STATEMENT,getHandleIdentifier()=XXX] (state=,code=0)

However, the MapReduce task that is submitted by the task to Yarn is successfully executed.

```
0: jdbc:hive2://189.120.204.104:21056/> select count(*) from test1;
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.psr.reducer=number;
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=number;
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reducers=number;
INFO : number of splits:1
INFO : Submitting tokens for job: job_1484563934624_0003
INFO : Kind: HDFS_DELEGATION_TOKEN, Service: ha-hdfs:haccluster, Ident: (HDFS_DELEGATION_TOKEN token 7 for admin)
INFO : Kind: HIVE_DELEGATION_TOKEN, Service: HiveServer2ImpersonationToken, Ident: 00 05 61 64 6d 69 6e 05 61 64 6d 69 6e 21 68 69 76 65 2f 68 61 64 6e 66 70 2e 68
85 ce e4 8a 01 59 ce 92 52 e4 8e 07 d8 0c
INFO : The url to track the job: https://189-120-204-104:26001/proxy/application_1484563934624_0003/
INFO : Starting Job = job_1484563934624_0003, Tracking URL = https://189-120-204-104:26001/proxy/application_1484563934624_0003/
INFO : Kill Command = /opt/*****/sigdata/FusionInsight-Hive-1.1.0/hadoop/bin/hadoop job -kill job_1484563934624_0003
INFO : Hadoop job information for Stage-1: number of mappers: 1 number of reducers: 1
INFO : 2017-01-17 11:46:12,579 Stage-1 map = 0%, reduce = 0%
INFO : 2017-01-17 11:46:23,243 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.32 sec
Error: Invalid OperationHandler: OperationHander [opType=EXECUTE_STATEMENT, getHandleIdentifier()=386323de-df1a-4299-826e-96368d4baf80] (state=,code=0)
0: jdbc:hive2://189.120.204.215:21056/>
```

## Cause Analysis

The cluster where the error occurs has two HiveServer instances. The error in the log of one HiveServer instance is the same as the error (Error: Invalid OperationHandler) reported on the client. In the log of the other HiveServer instance, **START\_UP** information similar to the following is printed when the error occurs, which indicates that the process is killed and restarted during that time. Because the HiveServer instance the task process plans to connect to is killed, it connects to the other healthy one, causing the error.

```
2017-02-15 14:40:11,309 | INFO | main | STARTUP_MSG:
/*****
STARTUP_MSG: Starting HiveServer2
STARTUP_MSG: host = XXX-120-85-154/XXX.120.85.154
STARTUP_MSG: args = []
STARTUP_MSG: version = 1.3.0
```

## Solution

Submit the task again and ensure that the HiveServer process is not manually restarted during task execution.

## 13.27 Error Message "Execution Error return code 2" Is Displayed When the SELECT Statement Is Executed

### Symptom

When the **select count(\*) from XXX;** statement is executed, the client reports the following error:

Error:Error while processing statement :FAILED:Execution Error,return code 2 from ...

**return code 2** indicates that the task fails because an error is reported during the execution of the MapReduce task.



```
00 jdbc:hive2://134.160.27.21:21066?select count(*) from src_gn_data_info_gz where day_id='18' and timenap='10';
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : in order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:496
INFO : Submitting tokens for job: job_1482323187492_57815
INFO : Kind: HDFS_DELEGATION_TOKEN, Service: ha-hdfs:hacluster, Ident: (HDFS_DELEGATION_TOKEN token 1083948 for boncusermm)
INFO : Kind: HIVE_DELEGATION_TOKEN, Service: HiveServer2ImpersonationToken, Ident: 00 0a 62 6f 6e 63 75 73 65 72 6d 6d 0a 62 6f 6e 63 75 73 65 72 6d 6d 21 68 05
74 55 0a 01 59 d4 b5 f8 55 8d 02 50 ae 0e 03 65
INFO : the url to track the job: https://hmcnc03-26901/proxy/application_1482323187492_57815/
INFO : Starting Job = job_1482323187492_57815, Tracking URL = https://hmcnc03-26901/proxy/application_1482323187492_57815/
INFO : Kill Command = /opt/BigData/FusionInsight_V100R002C68U10/FusionInsight-Hive-1.3.0/hive-1.3.0/bin/./../hadoop/bin/hadoop job -kill job_1482323187492_57815
INFO : Hadoop job information for Stage-1: number of mappers: 496; number of reducers: 1
INFO : 2017-01-18 16:21:00.906 Stage-1 map = 0%, reduce = 0%
INFO : 2017-01-18 16:21:10.957 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 50.53 sec
INFO : 2017-01-18 16:21:32.526 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 416.23 sec
INFO : 2017-01-18 16:21:35.035 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 1421.09 sec
INFO : 2017-01-18 16:21:36.331 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 2159.35 sec
INFO : 2017-01-18 16:21:37.810 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 2548.77 sec
INFO : 2017-01-18 16:21:39.126 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 3264.95 sec
INFO : 2017-01-18 16:21:40.509 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 3621.79 sec
INFO : 2017-01-18 16:21:41.710 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 3913.79 sec
INFO : 2017-01-18 16:21:42.899 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 4202.18 sec
INFO : 2017-01-18 16:21:44.037 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 4555.63 sec
INFO : 2017-01-18 16:21:45.119 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 4822.15 sec
INFO : 2017-01-18 16:21:46.213 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 5107.44 sec
INFO : 2017-01-18 16:21:47.380 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 5405.71 sec
INFO : 2017-01-18 16:21:48.407 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 5611.73 sec
INFO : 2017-01-18 16:21:49.483 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 5884.64 sec
INFO : 2017-01-18 16:21:50.565 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 5958.81 sec
INFO : 2017-01-18 16:21:51.641 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 6041.06 sec
INFO : 2017-01-18 16:21:52.745 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 6092.32 sec
INFO : 2017-01-18 16:22:08.525 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6078.4 sec
INFO : MapReduce Total cumulative CPU time: 0 days 1 hours 41 minutes 18 seconds 400 msec
ERROR : Ended Job = job_1482323187492_57815 with errors
Error: Error while processing statement: FAILED: Execution Error, return code 2 from org.apache.hadoop.hive.ql.exec.mr.MapReduceTask (state=08501.code=2)
00 jdbc:hive2://134.160.27.21:21066?>
```

## Cause Analysis

1. Go to the native Yarn page to check the MapReduce task logs. The check result shows that the error occurs due to an unidentified compression mode. The file name suffix is **.gzip** but the stack reports **.zlib**.

```
2017-01-18 16:22:07,566 INFO [main] org.apache.hadoop.hive.ql.exec.MapOperator: 4 Close done
2017-01-18 16:22:07,572 WARN [main] org.apache.hadoop.mapred.YarnChild: Exception running child : java.io.IOException: java.io.IOException: unknown compression method
at org.apache.hadoop.hive.io.HiveIOExceptionHandlerChain.handleRecordReaderNextException(HiveIOExceptionHandlerChain.java:121)
at org.apache.hadoop.hive.io.HiveIOExceptionHandlerChain.handleRecordReaderNextException(HiveIOExceptionHandlerChain.java:77)
at org.apache.hadoop.hive.ql.io.HiveContextAwareRecordReader.doNext(HiveContextAwareRecordReader.java:355)
at org.apache.hadoop.hive.ql.io.HiveRecordReader.doNext(HiveRecordReader.java:79)
at org.apache.hadoop.hive.ql.io.HiveRecordReader.doNext(HiveRecordReader.java:33)
at org.apache.hadoop.hive.ql.io.HiveContextAwareRecordReader.next(HiveContextAwareRecordReader.java:116)
at org.apache.hadoop.mapred.MapTask$TrackedRecordReader.moveToNext(MapTask.java:199)
at org.apache.hadoop.mapred.MapTask$TrackedRecordReader.next(MapTask.java:185)
at org.apache.hadoop.mapred.MapRunner.run(MapRunner.java:52)
at org.apache.hadoop.mapred.MapTask.runOldMapper(MapTask.java:453)
at org.apache.hadoop.mapred.MapTask.run(MapTask.java:343)
INFO : 2017-01-18 16:22:07,572 WARN [main] org.apache.hadoop.mapred.YarnChild: run(YarnChild.java:180)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1726)
at org.apache.hadoop.mapred.YarnChild.main(YarnChild.java:174)
Caused by: java.io.IOException: unknown compression method
at org.apache.hadoop.io.compress.zlib.ZlibDecompressor.inflateBytesDirect(Native Method)
at org.apache.hadoop.io.compress.zlib.ZlibDecompressor.decompress(ZlibDecompressor.java:226)
at org.apache.hadoop.io.compress.DecompressorStream.decompress(DecompressorStream.java:91)
at org.apache.hadoop.io.compress.DecompressorStream.read(DecompressorStream.java:85)
at java.io.InputStream.read(InputStream.java:101)
at org.apache.hadoop.util.LineReader.fillBuffer(LineReader.java:180)
at org.apache.hadoop.util.LineReader.readDefaultLine(LineReader.java:216)
at org.apache.hadoop.util.LineReader.readLine(LineReader.java:174)
at org.apache.hadoop.mapred.LineRecordReader.next(LineRecordReader.java:248)
at org.apache.hadoop.mapred.LineRecordReader.next(LineRecordReader.java:48)
at org.apache.hadoop.hive.ql.io.HiveContextAwareRecordReader.doNext(HiveContextAwareRecordReader.java:350)
... 13 more

2017-01-18 16:22:07,576 INFO [main] org.apache.hadoop.mapred.Task: Running cleanup for the task
```

2. Therefore, the HDFS file corresponding to the table that is queried may be incorrect. According to the file name printed in the Map log, download the file from HDFS to the local end. The file suffixed with **.gz** fails to be decompressed by running the **tar** command because its format is incorrect. Run the **file** command to check the file property. The command output shows that the file is compressed from the FAT system instead of UNIX.

```
[root@hndrn01 ~]# ls -l *.txt.gz
-rw-r--r-- 1 root root 101966463 Jan 18 20:13 201701180959589200740101.txt.gz
-rw-r--r-- 1 root root 90448283 Jan 18 19:55 20170118104000000740020.txt.gz
[root@hndrn01 ~]# file 201701181040000007400101.txt.gz
201701180959589200740101.txt.gz: gzip compressed data, was "201701180959589200740101.txt", from Unix, last modified: Wed Jan 18 09:59:52 2017
[root@hndrn01 ~]# file 20170118104000000740020.txt.gz
20170118104000000740020.txt.gz: gzip compressed data, from FAT filesystem (MS-DOS, OS/2, NT)
[root@hndrn01 ~]# tar -zxvf 20170118104000000740020.txt.gz
tar: This does not look like a tar archive
tar: Skipping to next header

gzip: stdin: decompression OK, trailing garbage ignored
tar: child returned status 2
tar: Error is not recoverable: exiting now
[root@hndrn01 ~]#
```

## Solution

Delete the file with an incorrect format from the HDFS directory or replace it with a correct one.

## 13.28 Failed to Perform drop partition When There Are a Large Number of Partitions

### Symptom

When the **drop partition** operation is performed, the following information is displayed:

```
MetaStoreClient lost connection. Attempting to reconnect. |
org.apache.hadoop.hive.metastore.RetryingMetaStoreClient.invoke(RetryingMetaStoreClient.java:187)
org.apache.thrift.transport.TTransportException
at org.apache.thrift.transport.TIOStreamTransport.read(TIOStreamTransport.java:132)
at org.apache.thrift.transport.TTransport.xxx(TTransport.java:86)
at org.apache.thrift.transport.TSaslTransport.readLength(TSaslTransport.java:376)
at org.apache.thrift.transport.TSaslTransport.readFrame(TSaslTransport.java:453)
at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:435)
...
```

As indicated by the MetaStore log, StackOverflow occurs.

```
2017-04-22 01:00:58,834 | ERROR | pool-6-thread-208 | java.lang.StackOverflowError
at org.datanucleus.store.rdbms.sql.SQLText.toSQL(SQLText.java:330)
at org.datanucleus.store.rdbms.sql.SQLText.toSQL(SQLText.java:339)
at org.datanucleus.store.rdbms.sql.SQLText.toSQL(SQLText.java:339)
at org.datanucleus.store.rdbms.sql.SQLText.toSQL(SQLText.java:339)
at org.datanucleus.store.rdbms.sql.SQLText.toSQL(SQLText.java:339)
```

### Cause Analysis

The processing logic of the drop partition operation is to find all the partitions that meet the conditions, combine them, and delete them together. However, because the number of partitions is too large and the data stack for deleting metadata is deep, StackOverflow errors occur.

### Solution

Delete partitions in batches.

## 13.29 Failed to Start the Local Task When the Join Operation Is Performed

### Symptom

1. When operations such as JOIN are performed for a small amount of data, a local task will be started. However, the execution fails and reports the following error:

```
jdbc:hive2://10.*.*:21066/> select a.name ,b.sex from student a join student1 b on (a.name = b.name);
ERROR : Execution failed with exit status: 1
ERROR : Obtaining error information
ERROR :
Task failed!
Task ID:
Stage-4
...
Error: Error while processing statement: FAILED: Execution Error, return code 1 from
```

- ```
org.apache.hadoop.hive ql.exec.mr.MapredLocalTask (state=08S01,code=1)
...
2. The HiveServer log shows that the local task fails to start.
2018-04-25 16:37:19,296 | ERROR | HiveServer2-Background-Pool: Thread-79 | Execution failed with
exit status: 1 | org.apache.hadoop.hive ql.session.SessionState
$LogHelper.printError(SessionState.java:1016)
2018-04-25 16:37:19,296 | ERROR | HiveServer2-Background-Pool: Thread-79 | Obtaining error
information | org.apache.hadoop.hive ql.session.SessionState
$LogHelper.printError(SessionState.java:1016)
2018-04-25 16:37:19,297 | ERROR | HiveServer2-Background-Pool: Thread-79 |
Task failed!
Task ID:
  Stage-4
Logs:
  | org.apache.hadoop.hive ql.session.SessionState$LogHelper.printError(SessionState.java:1016)
2018-04-25 16:37:19,297 | ERROR | HiveServer2-Background-Pool: Thread-79 | /var/log/Bigdata/hive/
hiveserver/hive.log | org.apache.hadoop.hive ql.session.SessionState
$LogHelper.printError(SessionState.java:1016)
2018-04-25 16:37:19,297 | ERROR | HiveServer2-Background-Pool: Thread-79 | Execution failed with
exit status: 1 |
org.apache.hadoop.hive ql.exec.mr.MapredLocalTask.executeInChildVM(MapredLocalTask.java:342)
2018-04-25 16:37:19,309 | ERROR | HiveServer2-Background-Pool: Thread-79 | FAILED: Execution
Error, return code 1 from org.apache.hadoop.hive ql.exec.mr.MapredLocalTask |
org.apache.hadoop.hive ql.session.SessionState$LogHelper.printError(SessionState.java:1016)
...
2018-04-25 16:37:36,438 | ERROR | HiveServer2-Background-Pool: Thread-88 | Error running hive
query: | org.apache.hive.service.cli.operation.SQLOperation$1$1.run(SQLOperation.java:248)
org.apache.hive.service.cli.HiveSQLException: Error while processing statement: FAILED: Execution
Error, return code 1 from org.apache.hadoop.hive ql.exec.mr.MapredLocalTask
  at org.apache.hive.service.cli.operation.Operation.toSQLException(Operation.java:339)
  at org.apache.hive.service.cli.operation.SQLOperation.runQuery(SQLOperation.java:169)
  at org.apache.hive.service.cli.operation.SQLOperation.access$200(SQLOperation.java:75)
  at org.apache.hive.service.cli.operation.SQLOperation$1$1.run(SQLOperation.java:245)
  at java.security.AccessController.doPrivileged(Native Method)
  at javax.security.auth.Subject.doAs(Subject.java:422)
  at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1710)
  at org.apache.hive.service.cli.operation.SQLOperation$1.run(SQLOperation.java:258)
  at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
  at java.util.concurrent.FutureTask.run(FutureTask.java:266)
  at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
  at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
  at java.lang.Thread.run(Thread.java:745)
...
3. The hs_err_pid_****.log file in the HiveServer log directory /var/log/Bigdata/hive/hiveserver contains an error about insufficient memory.
# There is insufficient memory for the Java Runtime Environment to continue.
# Native memory allocation (mmap) failed to map 20776943616 bytes for committing reserved
memory.
...
```

## Cause Analysis

When Hive executes JOIN for a small amount of data, MapJoin is generated. During MapJoin execution, a local task is started. JVM memory launched by the local task inherits the memory of the parent process.

When multiple JOIN operations are executed, multiple local tasks are started. If the host is out of memory, the local tasks fail to start.

## Solution

**Step 1** Go to the Hive configuration page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console, choose **Components** > **Hive** > **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 3.x or later: Log in to FusionInsight Manager and choose **Cluster**. Click the name of the target cluster, and choose **Services > Hive > Configurations > All Configurations**.

**Step 2** Search for the **hive.auto.convert.join** parameter and change the value of **hive.auto.convert.join** in Hive to **false**. Save the configuration and restart the service.

The value change may deteriorate service performance. You can perform the next step to avoid adverse impacts on the performance.

**Step 3** Search for the **HIVE\_GC\_OPTS** parameter and decrease the value of **Xms** based on service requirements. The minimum value is half that of **Xmx**. After the modification, save the configuration and restart the service.

----End

## 13.30 WebHCat Fails to Be Started After the Hostname Is Changed

### Symptom

WebHCat fails to be started after the hostname is changed.

The following error is reported in the WebHCat startup log (`/var/log/Bigdata/hive/webhcat/hive.log`) of the corresponding node.

```
org.apache.hadoop.security.authentication.client.AuthenticationException: GSSException: No valid credentials provided (Mechanism level: Server not found in Kerberos database (7))
at org.apache.hadoop.hive.cm.utils.WebHCatAuthenticator.doSpnegoSequence(WebHCatAuthenticator.java:302)
at org.apache.hadoop.hive.cm.utils.WebHCatAuthenticator.authenticate(WebHCatAuthenticator.java:149)
at org.apache.hadoop.hive.cm.monitor.WebHCatHealthChecker.renewToken(WebHCatHealthChecker.java:186)
at org.apache.hadoop.hive.cm.monitor.WebHCatHealthChecker.checkWebHCat(WebHCatHealthChecker.java:119)
at org.apache.hadoop.hive.cm.monitor.WebHCatHealthChecker.run(WebHCatHealthChecker.java:168)
at java.lang.Thread.run(Thread.java:745)
Caused by: GSSException: No valid credentials provided (Mechanism level: Server not found in Kerberos database (7)) - UNKNOWN_SERVER
at sun.security.jgss.krb5.Krb5Context.initSecContext(Krb5Context.java:990)
at sun.security.jgss.GSSContextImpl.initSecContext(GSSContextImpl.java:248)
at sun.security.jgss.GSSContextImpl.initSecContext(GSSContextImpl.java:179)
at org.apache.hadoop.hive.cm.utils.WebHCatAuthenticator$1.run(WebHCatAuthenticator.java:277)
at org.apache.hadoop.hive.cm.utils.WebHCatAuthenticator$1.run(WebHCatAuthenticator.java:253)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.hive.cm.utils.WebHCatAuthenticator.doSpnegoSequence(WebHCatAuthenticator.java:253)
... 5 more
Caused by: KrbException: Server not found in Kerberos database (7) - UNKNOWN_SERVER
at sun.security.krb5.KrbTgsRep.<init>(KrbTgsRep.java:73)
at sun.security.krb5.KrbTgsReq.getReply(KrbTgsReq.java:251)
at sun.security.krb5.KrbTgsReq.sendAndGetCreds(KrbTgsReq.java:262)
at sun.security.krb5.internal.CredentialsUtil.getServiceCreds(CredentialsUtil.java:308)
at sun.security.krb5.internal.CredentialsUtil.acquireServiceCreds(CredentialsUtil.java:126)
at sun.security.krb5.Credentials.acquireServiceCreds(Credentials.java:458)
at sun.security.jgss.krb5.Krb5Context.initSecContext(Krb5Context.java:693)
... 12 more
Caused by: KrbException: Identifier doesn't match expected value (906)
at sun.security.krb5.internal.KrbRep.init(KrbRep.java:140)
at sun.security.krb5.internal.TGSRep.init(TGSRep.java:65)
at sun.security.krb5.internal.TGSRep.<init>(TGSRep.java:60)
at sun.security.krb5.KrbTgsRep.<init>(KrbTgsRep.java:55)
```

### Cause Analysis

1. The server account of the MRS WebHCat role contains the hostname. If you change the hostname after the cluster installation, WebHCat fails to start.
2. The one-to-many or many-to-one association between IP addresses and hostnames is configured in the `/etc/hosts` file. As a result, the IP address and hostname cannot be obtained correctly after the **hostname** and **hostname -i** commands are executed.

## Solution

- Step 1** Change the hostname of the modified node to the initial hostname of the cluster.
  - Step 2** Check whether the `/etc/hosts` of the node where WebHCat is located is correctly configured.
  - Step 3** Restart WebHCat.
- End

## 13.31 An Error Is Reported When the Hive Sample Program Is Running After the Domain Name of a Cluster Is Changed

### Symptom

In the sample code for Hive secondary development, error "No rules applied to \*\*\*\*" is reported.

```
java.io.IOException: Login failure for platformUser@ADHOC.COM from keytab user.keytab: javax.security.auth.login.LoginException: java.lang.IllegalArgumentException: Illegal principal name platformUser@ADHOC.COM: org.apache.hadoop.security.authentication.util.KerberosName$NoMatchingRule: No rules applied to platformUser@ADHOC.COM
    at org.apache.hadoop.security.UserGroupInformation.loginUserFromKeytab(LoginUtil.java:979)
    at com.huawei.adhoc.connector.factory.LoginUtil.loginHadoop(LoginUtil.java:311)
    at com.huawei.adhoc.connector.factory.LoginUtil.login(LoginUtil.java:134)
    at com.huawei.adhoc.connector.factory.C70ConnectorFactory.getConnection(C70ConnectorFactory.java:92)
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:498)
    at com.huawei.adhoc.jdbc.connection.util.GetConnectionHolder70.run(ConnectionUtil.java:238)
    at java.lang.Thread.run(Thread.java:745)
Caused by: javax.security.auth.login.LoginException: java.lang.IllegalArgumentException: Illegal principal name platformUser@ADHOC.COM: org.apache.hadoop.security.authentication.util.KerberosName$NoMatchingRule: No rules applied to platformUser@ADHOC.COM
    at org.apache.hadoop.security.UserGroupInformation$HadoopLoginModule.commit(UserGroupInformation.java:202)
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:498)
    at javax.security.auth.login.LoginContext.invoke(LoginContext.java:755)
    at javax.security.auth.login.LoginContext.access$000(LoginContext.java:195)
```

### Cause Analysis

- The sample code for Hive secondary development loads `core-site.xml` file that is loaded through classload by default. Therefore, you need to put the configuration file to the `classpath` directory of the startup program.
- If the domain name of the cluster is changed, the `core-site.xml` file will change. You need to download the latest `core-site.xml` file and save it to the `classpath` directory where the sample code for Hive secondary development is located.

## Solution

- Step 1** Download the latest client of the Hive cluster to obtain the latest `core-site.xml` file.
  - Step 2** Save the `core-site.xml` file to the `classpath` directory where the sample code process for Hive secondary development is located.
- End

## 13.32 Hive MetaStore Exception Occurs When the Number of DBService Connections Exceeds the Upper Limit

### Symptom

By default, the maximum number of connections to DBService is 300. If the number of connections is greater than 300 due to heavy traffic, an exception occurs in MetaStore and error "slots are reserved for non-replication superuser connections" is reported.

```
2018-04-26 14:58:55,657 | ERROR | BoneCP-pool-watch-thread | Failed to acquire connection to
jdbc:postgresql://10.**.*:20051/hivemeta?socketTimeout=60. Sleeping for 1000 ms. Attempts left: 9 |
com.jolbox.bonecp.BoneCP.obtainInternalConnection(BoneCP.java:292)
org.postgresql.util.PSQLException: FATAL: remaining connection slots are reserved for non-replication
superuser connections
    at org.postgresql.core.v3.ConnectionFactoryImpl.readStartupMessages(ConnectionFactoryImpl.java:643)
    at org.postgresql.core.v3.ConnectionFactoryImpl.openConnectionImpl(ConnectionFactoryImpl.java:184)
    at org.postgresql.core.ConnectionFactory.openConnection(ConnectionFactory.java:64)
    at org.postgresql.jdbc2.AbstractJdbc2Connection.<init>(AbstractJdbc2Connection.java:124)
    at org.postgresql.jdbc3.AbstractJdbc3Connection.<init>(AbstractJdbc3Connection.java:28)
    at org.postgresql.jdbc3g.AbstractJdbc3gConnection.<init>(AbstractJdbc3gConnection.java:20)
    at org.postgresql.jdbc4.AbstractJdbc4Connection.<init>(AbstractJdbc4Connection.java:30)
    at org.postgresql.jdbc4.Jdbc4Connection.<init>(Jdbc4Connection.java:22)
    at org.postgresql.Driver.makeConnection(Driver.java:392)
    at org.postgresql.Driver.connect(Driver.java:266)
    at java.sql.DriverManager.getConnection(DriverManager.java:664)
    at java.sql.DriverManager.getConnection(DriverManager.java:208)
    at com.jolbox.bonecp.BoneCP.obtainRawInternalConnection(BoneCP.java:361)
    at com.jolbox.bonecp.BoneCP.obtainInternalConnection(BoneCP.java:269)
    at com.jolbox.bonecp.ConnectionHandle.<init>(ConnectionHandle.java:242)
    at com.jolbox.bonecp.PoolWatchThread.fillConnections(PoolWatchThread.java:115)
    at com.jolbox.bonecp.PoolWatchThread.run(PoolWatchThread.java:82)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
    at java.lang.Thread.run(Thread.java:745)
```

### Cause Analysis

Heavy service traffic causes more than 300 connections to DBService, and the maximum number of connections to DBService needs to be increased.

### Solution

**Step 1** Go to the DBService configuration page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console, choose **Components** > **DBService** > **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 3.x or later: Log in to FusionInsight Manager and choose **Cluster** > **Services** > **DBService** > **Configurations** > **All Configurations**.

**Step 2** Search for **dbservice.database.max.connections** and set it to a proper value not greater than **1000**.

**Step 3** Save the configuration and restart the affected services or instances.

**Step 4** If the fault persists, check the service code for any connection leaks.

----End

## 13.33 Error Message "Failed to execute session hooks: over max connections" Is Displayed on the Beeline Client

### Symptom

The default maximum connections to HiveServer are 200. When the number of connections exceeds 200, Beeline reports error "Failed to execute session hooks: over max connections:".

```
beeline> [root@172-27-16-38 c70client]# beeline
Connecting to
jdbc:hive2://129.188.82.38:24002,129.188.82.36:24002,129.188.82.35:24002/?serviceDiscoveryMode=zooKeeper;
zooKeeperNamespace=hiveserver2;sasl.qop=auth-conf;auth=KERBEROS;principal=hive/
hadoop.hadoop.com@HADOOP.COM
Debug is true storeKey false useTicketCache true useKeyTab false doNotPrompt false ticketCache is null
isInitiator true KeyTab is null refreshKrb5Config is false principal is null tryFirstPass is false useFirstPass is
false storePass is false clearPass is false
Acquire TGT from Cache
Principal is xxx@HADOOP.COM
Commit Succeeded

Error: Failed to execute session hooks: over max connections. (state=,code=0)
Beeline version 1.2.1 by Apache Hive
```

The HiveServer log (**/var/log/Bigdata/hive/hiveserver/hive.log**) shows that error "over max connections" is reported.

```
2018-05-03 04:31:56,728 | WARN | HiveServer2-Handler-Pool: Thread-137 | Error opening session: |
org.apache.hive.service.cli.thrift.ThriftCLIService.OpenSession(ThriftCLIService.java:542)
org.apache.hive.service.cli.HiveSQLException: Failed to execute session hooks: over max connections.
    at org.apache.hive.service.cli.session.SessionManager.openSession(SessionManager.java:322)
    at org.apache.hive.service.cli.CLIService.openSessionWithImpersonation(CLIService.java:189)
    at org.apache.hive.service.cli.thrift.ThriftCLIService.getSessionHandle(ThriftCLIService.java:663)
    at org.apache.hive.service.cli.thrift.ThriftCLIService.OpenSession(ThriftCLIService.java:527)
    at org.apache.hive.service.cli.thrift.TCLIService$Processor$OpenSession.getResult(TCLIService.java:1257)
    at org.apache.hive.service.cli.thrift.TCLIService$Processor$OpenSession.getResult(TCLIService.java:1242)
    at org.apache.thrift.ProcessFunction.process(ProcessFunction.java:39)
    at org.apache.thrift.TBaseProcessor.process(TBaseProcessor.java:39)
    at org.apache.hadoop.hive.thrift.HadoopThriftAuthBridge$Server
    $TUGIAssumingProcessor.process(HadoopThriftAuthBridge.java:710)
    at org.apache.thrift.server.TThreadPoolServer$WorkerProcess.run(TThreadPoolServer.java:286)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
    at java.lang.Thread.run(Thread.java:745)
Caused by: org.apache.hive.service.cli.HiveSQLException: over max connections.
    at
    org.apache.hadoop.hive.transporthook.SessionControllerTsasTransportHook.checkTotalSessionNumber(Sessi
    onControllerTsasTransportHook.java:208)
    at
    org.apache.hadoop.hive.transporthook.SessionControllerTsasTransportHook.postOpen(SessionControllerTsas
    TransportHook.java:163)
```

```
at
org.apache.hadoop.hive.transporthook.SessionControllerTsslTransportHook.run(SessionControllerTsslTransportHook.java:134)
at org.apache.hive.service.cli.session.SessionManager.executeSessionHooks(SessionManager.java:432)
at org.apache.hive.service.cli.session.SessionManager.openSession(SessionManager.java:314)
... 12 more
```

## Cause Analysis

Heavy service traffic causes the number of connections to one HiveServer node to exceed 200, and the maximum number of connections to HiveServer needs to be increased.

## Solution

**Step 1** Go to the Hive configuration page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console, choose **Components** > **Hive** > **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 3.x or later: Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Hive** > **Configurations** > **All Configurations**.

**Step 2** Search for **hive.server.session.control.maxconnections** and set it to a proper value not greater than **1000**.

**Step 3** Save the configuration and restart the affected services or instances.

----End

## 13.34 Error Message "OutOfMemoryError" Is Displayed on the Beeline Client

### Symptom

When a large amount of data is queried on the Beeline client, the message "OutOFMemoryError: Java heap space" is displayed. The detailed error information is as follows:

```
org.apache.thrift.TException: Error in calling method FetchResults
at org.apache.hive.jdbc.HiveConnection$SynchronizedHandler.invoke(HiveConnection.java:1514)
at com.sun.proxy.$Proxy4.FetchResults(Unknown Source)
at org.apache.hive.jdbc.HiveQueryResultSet.next(HiveQueryResultSet.java:358)
at org.apache.hive.beeline.BufferedRows.<init>(BufferedRows.java:42)
at org.apache.hive.beeline.BeeLine.print(BeeLine.java:1856)
at org.apache.hive.beeline.Commands.execute(Commands.java:873)
at org.apache.hive.beeline.Commands.sql(Commands.java:714)
at org.apache.hive.beeline.BeeLine.dispatch(BeeLine.java:1035)
at org.apache.hive.beeline.BeeLine.execute(BeeLine.java:821)
at org.apache.hive.beeline.BeeLine.begin(BeeLine.java:778)
at org.apache.hive.beeline.BeeLine.mainWithInputRedirection(BeeLine.java:486)
at org.apache.hive.beeline.BeeLine.main(BeeLine.java:469)
```



```
Caused by: java.lang.OutOfMemoryError: Java heap space
  at com.sun.crypto.provider.CipherCore.doFinal(CipherCore.java:959)
  at com.sun.crypto.provider.CipherCore.doFinal(CipherCore.java:824)
  at com.sun.crypto.provider.AESCipher.engineDoFinal(AESCipher.java:436)
  at javax.crypto.Cipher.doFinal(Cipher.java:2223)
  at sun.security.krb5.internal.crypto.dk.AesDkCrypto.decryptCTS(AesDkCrypto.java:414)
  at sun.security.krb5.internal.crypto.dk.AesDkCrypto.decryptRaw(AesDkCrypto.java:291)
  at sun.security.krb5.internal.crypto.Aes256.decryptRaw(Aes256.java:86)
  at sun.security.jgss.krb5.CipherHelper.aes256Decrypt(CipherHelper.java:1397)
  at sun.security.jgss.krb5.CipherHelper.decryptData(CipherHelper.java:576)
  at sun.security.jgss.krb5.WrapToken_v2.getData(WrapToken_v2.java:130)
  at sun.security.jgss.krb5.WrapToken_v2.getData(WrapToken_v2.java:105)
  at sun.security.jgss.krb5.Krb5Context.unwrap(Krb5Context.java:1058)
  at sun.security.jgss.GSSContextImpl.unwrap(GSSContextImpl.java:403)
  at com.sun.security.sasl.gsskerb.GssKrb5Base.unwrap(GssKrb5Base.java:77)
  at org.apache.thrift.transport.TSaslTransport$SaslParticipant.unwrap(TSaslTransport.java:559)
  at org.apache.thrift.transport.TSaslTransport.readFrame(TSaslTransport.java:462)
  at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:435)
  at org.apache.thrift.transport.TSaslClientTransport.read(TSaslClientTransport.java:37)
  at org.apache.thrift.transport.TTransport.xxx(TTransport.java:86)
  at org.apache.hadoop.hive.thrift.TFilterTransport.xxx(TFilterTransport.java:62)
  at org.apache.thrift.protocol.TBinaryProtocol.xxx(TBinaryProtocol.java:429)
  at org.apache.thrift.protocol.TBinaryProtocol.readI32(TBinaryProtocol.java:318)
  at org.apache.thrift.protocol.TBinaryProtocol.readMessageBegin(TBinaryProtocol.java:219)
  at org.apache.thrift.TServiceClient.receiveBase(TServiceClient.java:77)
  at org.apache.hive.service.cli.thrift.TCLIService$Client.recv_FetchResults(TCLIService.java:505)
  at org.apache.hive.service.cli.thrift.TCLIService$Client.FetchResults(TCLIService.java:492)
  at sun.reflect.GeneratedMethodAccessor2.invoke(Unknown Source)
  at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
  at java.lang.reflect.Method.invoke(Method.java:498)
  at org.apache.hive.jdbc.HiveConnection$SynchronizedHandler.invoke(HiveConnection.java:1506)
  at com.sun.proxy.$Proxy4.FetchResults(Unknown Source)
  at org.apache.hive.jdbc.HiveQueryResultSet.next(HiveQueryResultSet.java:358)
Error: Error retrieving next row (state=,code=0)
```

## Cause Analysis

- The data volume is excessively large.
- Users use the **select \* from table\_name;** statement for query in the whole table. There is a large amount of data in the table.
- The default startup memory of Beeline is 128 MB. The returned result set is too large during query, overloading Beeline.

## Solution

- Step 1** Before running **select count(\*) from table\_name;**, check the amount of data to be queried and determine whether to display data of this magnitude in Beeline.
- Step 2** If a certain amount of data needs to be displayed, adjust the JVM parameter of the Hive client. Add **export HIVE\_OPTS=-Xmx1024M** (change the value based on service requirements) to **component\_env** in the **/Hive** directory of the Hive client. Run the **source Client directory/bigdata\_env** command again to configure environment variables.

----End

## 13.35 Task Execution Fails Because the Input File Number Exceeds the Threshold

### Symptom

When Hive performs a query operation, error message "Job Submission failed with exception 'java.lang.RuntimeException(input file number exceeded the limits in the conf;input file num is: 2380435,max heap memory is: 16892035072,the limit conf is: 500000/4)'" is displayed. The value in the error message varies depending on the actual situation. The error details are as follows:

```
ERROR : Job Submission failed with exception 'java.lang.RuntimeException(input file numbers exceeded the limits in the conf;
input file num is: 2380435 ,
max heap memory is: 16892035072 ,
the limit conf is: 500000/4)'
java.lang.RuntimeException: input file numbers exceeded the limits in the conf;
input file num is: 2380435 ,
max heap memory is: 16892035072 ,
the limit conf is: 500000/4
    at org.apache.hadoop.hive ql.exec.mr.ExecDriver.checkFileNum(ExecDriver.java:545)
    at org.apache.hadoop.hive ql.exec.mr.ExecDriver.execute(ExecDriver.java:430)
    at org.apache.hadoop.hive ql.exec.mr.MapRedTask.execute(MapRedTask.java:137)
    at org.apache.hadoop.hive ql.exec.Task.executeTask(Task.java:158)
    at org.apache.hadoop.hive ql.exec.TaskRunner.runSequential(TaskRunner.java:101)
    at org.apache.hadoop.hive ql.Driver.launchTask(Driver.java:1965)
    at org.apache.hadoop.hive ql.Driver.execute(Driver.java:1723)
    at org.apache.hadoop.hive ql.Driver.runInternal(Driver.java:1475)
    at org.apache.hadoop.hive ql.Driver.run(Driver.java:1283)
    at org.apache.hadoop.hive ql.Driver.run(Driver.java:1278)
    at org.apache.hive.service.cli.operation.SQLOperation.runQuery(SQLOperation.java:167)
    at org.apache.hive.service.cli.operation.SQLOperation.access$200(SQLOperation.java:75)
    at org.apache.hive.service.cli.operation.SQLOperation$1$1.run(SQLOperation.java:245)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:422)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1710)
    at org.apache.hive.service.cli.operation.SQLOperation$1.run(SQLOperation.java:258)
    at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
    at java.util.concurrent.FutureTask.run(FutureTask.java:266)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
    at java.lang.Thread.run(Thread.java:745)
```

Error: Error while processing statement: FAILED: Execution Error, return code 1 from org.apache.hadoop.hive ql.exec.mr.MapRedTask (state=08S01,code=1)

### Cause Analysis

MRS uses the ratio of maximum files to the maximum HiveServer heap memory to determine the number of input files allowed in a MapReduce job submission. Default value **500000/4** indicates that each 4 GB of heap memory allows a maximum of 500,000 input files. An error occurs if the number of input files exceeds this limit.

### Solution

**Step 1** Go to the Hive configuration page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console, choose **Components** > **Hive** > **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 3.x or later: Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Hive** > **Configurations** > **All Configurations**.

**Step 2** Search for **hive.mapreduce.input.files2memory** and set it to a proper value based on the actual memory and task.

**Step 3** Save the configuration and restart the affected services or instances.

**Step 4** If the fault persists, adjust the GC parameter of the HiveServer based on service requirements.

----End

## 13.36 Hive Task Execution Fails Because of Stack Memory Overflow

### Symptom

When Hive performs a query operation, error "Error running child: java.lang.StackOverflowError" is reported. The error details are as follows:

```
FATAL [main] org.apache.hadoop.mapred.YarnChild: Error running child : java.lang.StackOverflowError
at org.apache.hive.com.esotericsoftware.kryo.io.Input.readVarInt(Input.java:355)
at
org.apache.hive.com.esotericsoftware.kryo.util.DefaultClassResolver.readName(DefaultClassResolver.java:127)
at
org.apache.hive.com.esotericsoftware.kryo.util.DefaultClassResolver.readClass(DefaultClassResolver.java:115)
at org.apache.hive.com.esotericsoftware.kryo.Kryo.readClass(Kryo.java:656)
at org.apache.hive.com.esotericsoftware.kryo.kryo.readClassAndObject(Kryo.java:767)
at
org.apache.hive.com.esotericsoftware.kryo.serializers.collectionSerializer.read(CollectionSerializer.java:112)
```

```
2018-08-07 09:16:54,243 INFO [main] org.apache.hadoop.hive.ql.exec.Utilities: PLAN PATH = hdfs://hacluster/tmp/hive-scratch/lzy/dc3f8015-1b1e-4234-b45e-3f919fcaa485/hive_2018-08-07_09-13-50
676_7895353416339631598-383269/-mr-10804/3514ec7f-5268-4431-9c17-f2814f5f99b7/map.xml
2018-08-07 09:16:54,243 INFO [main] org.apache.hadoop.hive.ql.exec.Utilities: *****non-local mode*****
2018-08-07 09:16:54,243 INFO [main] org.apache.hadoop.hive.ql.exec.Utilities: local path = hdfs://hacluster/tmp/hive-scratch/lzy/dc3f8015-1b1e-4234-b45e-3f919fcaa485/hive_2018-08-07_09-13-5
0_676_7895353416339631598-383269/-mr-10804/3514ec7f-5268-4431-9c17-f2814f5f99b7/map.xml
2018-08-07 09:16:54,244 INFO [main] org.apache.hadoop.hive.ql.exec.Utilities: Open file to read in plan: hdfs://hacluster/tmp/hive-scratch/lzy/dc3f8015-1b1e-4234-b45e-3f919fcaa485/hive_2018
-08-07_09-13-50_676_7895353416339631598-383269/-mr-10804/3514ec7f-5268-4431-9c17-f2814f5f99b7/map.xml
2018-08-07 09:16:54,260 INFO [main] org.apache.hadoop.hive.ql.log.PerfLogger: <PERFLOG method=deserializePlan from=org.apache.hadoop.hive.ql.exec.Utilities>
2018-08-07 09:16:54,260 INFO [main] org.apache.hadoop.hive.ql.exec.Utilities: Deserializing MapWork via kryo
2018-08-07 09:16:54,468 FATAL [main] org.apache.hadoop.mapred.YarnChild: Error running child : java.lang.StackOverflowError
at org.apache.hive.com.esotericsoftware.kryo.io.Input.readVarInt(Input.java:355)
at org.apache.hive.com.esotericsoftware.kryo.util.DefaultClassResolver.readName(DefaultClassResolver.java:127)
at org.apache.hive.com.esotericsoftware.kryo.util.DefaultClassResolver.readClass(DefaultClassResolver.java:115)
at org.apache.hive.com.esotericsoftware.kryo.Kryo.readClass(Kryo.java:656)
at org.apache.hive.com.esotericsoftware.kryo.kryo.readClassAndObject(Kryo.java:767)
at org.apache.hive.com.esotericsoftware.kryo.serializers.collectionSerializer.read(CollectionSerializer.java:112)
```

3193,1-0 50%

### Cause Analysis

Error "java.lang.StackOverflowError" indicates the memory overflow of the thread stack. It may occur if there are multiple levels of calls (for example, infinite recursive calls) or the thread stack is too small.

## Solution

Adjust the stack memory in the JVM parameters of the Map and Reduce stages during execution of a MapReduce job, that is, **mapreduce.map.java.opts** (adjusting the stack memory of Map) and **mapreduce.reduce.java.opts** (adjusting the stack memory of Reduce). The following uses the **mapreduce.map.java.opts** parameter as an example.

- To increase the Map memory temporarily (only valid for Beeline):  
Run the **set mapreduce.map.java.opts=-Xss8G;** command on the Beeline client. (Change the value as required.)
- To permanently increase the Map memory specified by the **mapreduce.map.memory.mb** and **mapreduce.map.java.opts** parameters:
  - a. Go to the Hive configuration page.
    - For versions earlier than MRS 3.x: Click the cluster name on the MRS console, choose **Components > Hive > 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 3.x or later: Log in to FusionInsight Manager and choose **Cluster**. Click the name of the target cluster and choose **Services > Hive > Configurations > All Configurations**.
- b. Add custom parameter **mapreduce.map.java.opts** and set it to a proper value.
- c. Save the configuration and restart the affected services or instances.  
Note that the modification takes effect after a service restart. During the restart, the Hive service is unavailable.

## 13.37 Task Failed Due to Concurrent Writes to One Table or Partition

### Symptom

When Hive executes an INSERT statement, an error is reported indicating that a file or directory already exists or is cleared in HDFS. The error details are as follows:

```
2019-03-18 14:34:23.016 | WARN | HiveServer2-Background-Pool: Thread-1179606 | Failed to move to trash: hdfs://hacluster/user/hive/warehouse/frpdb.db/dw_fixed_cost_xm_temp5_f/000000_0; Force to delete it. | org.apache.hadoop.hive.common.FileUtils.moveToTrash(FileUtils.java:651)
2019-03-18 14:34:23.017 | INFO | HiveServer2-Background-Pool: Thread-1179604 | Moved to trash: hdfs://hacluster/user/hive/warehouse/frpdb.db/dw_fixed_cost_xm_temp6_f/000000_0 | org.apache.hadoop.hive.common.FileUtils.moveToTrash(FileUtils.java:644)
2019-03-18 14:34:23.017 | ERROR | HiveServer2-Background-Pool: Thread-1179606 | Failed to delete hdfs://hacluster/user/hive/warehouse/frpdb.db/dw_fixed_cost_xm_temp5_f/000000_0 | org.apache.hadoop.hive.common.FileUtils.moveToTrash(FileUtils.java:660)
2019-03-18 14:34:23.017 | ERROR | HiveServer2-Background-Pool: Thread-1179606 | Failed with exception Destination directory hdfs://hacluster/user/hive/warehouse/frpdb.db/dw_fixed_cost_xm_temp5_f has not been cleaned up.
org.apache.hadoop.hive.ql.metadata.HiveException: Destination directory hdfs://hacluster/user/hive/warehouse/frpdb.db/dw_fixed_cost_xm_temp5_f has not been cleaned up.
at org.apache.hadoop.hive.ql.metadata.Hive.replaceFile(Hive.java:2974)
at org.apache.hadoop.hive.ql.metadata.Hive.loadTable(Hive.java:1664)
at org.apache.hadoop.hive.ql.exec.MoveTask.execute(MoveTask.java:374)
at org.apache.hadoop.hive.ql.exec.Task.executeTask(Task.java:158)
at org.apache.hadoop.hive.ql.exec.TaskRunner.run(TaskRunner.java:101)
```

### Cause Analysis

1. Check the start time and end time of the task based on the HiveServer audit logs.

2. Check whether data is inserted into the same table or partition in the time segment.
3. Hive does not support concurrent data insertion for a table or partition. As a result, multiple tasks perform operations on the same temporary data directory, and one task moves the data of another task, causing task failure.

## Solution

The service logic is modified so that data is inserted to the same table or partition in single thread mode.

## 13.38 Hive Task Failed Due to a Lack of HDFS Directory Permission

### Symptom

An error message is displayed, indicating that the user does not have the permission to access the HDFS directory:

```
2019-04-09 17:49:19,845 | ERROR | HiveServer2-Background-Pool: Thread-3160445 | Job Submission failed with exception 'org.apache.hadoop.security.AccessControlException(Permission denied: user=hive_quanxian, access=READ_EXECUTE, inode="/user/hive/warehouse/bigdata.db/gd_ga_wa_swryswj":zhongao:hive:drwx-----
at
org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkAccessAcl(FSPermissionChecker.java:426)
)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:329)
at
org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSubAccess(FSPermissionChecker.java:300)
at
org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:241)
at
com.xxx.hadoop.adapter.hdfs.plugin.HWAccessControlEnforce.checkPermission(HWAccessControlEnforce.java:69)
at
org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:190)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1910)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1894)
at
org.apache.hadoop.hdfs.server.namenode.FSDirStatAndListingOp.getContentSummary(FSDirStatAndListingOp.java:135)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getContentSummary(FSNamesystem.java:3983)
at
org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getContentSummary(NameNodeRpcServer.java:1342)
at
org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getContentSummary(ClientNamenodeProtocolServerSideTranslatorPB.java:925)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:616)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:973)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2260)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2256)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1781)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2254)
)'
```

## Cause Analysis

1. According to the stack information, the permission on the subdirectory fails to be checked.  
org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSubAccess(FSPermissionChecker.java:300)
2. Check the permission of all files and directories in HDFS. The permission of a directory is 700 (only the file owner can access the directory). It is confirmed that an abnormal directory exists.

```
[root@hdfsServer ~]# find /user/hive/warehouse/bigdata.db/gd_ga_wa_wrysj/l/.hive-staging_hive_2019-03-16_08-10-08_139_4871120769486435512-104876 -type d -exec ls -ld {} \;
Found 2 items
drwxr-xr-x 1 hive 9 2019-03-16 08:22 /user/hive/warehouse/bigdata.db/gd_ga_wa_wrysj/l/.hive-staging_hive_2019-03-16_08-10-08_139_4871120769486435512-104876/_task_tmp_-ext-10000
drwxr-xr-x 1 hive 8 2019-03-16 08:22 /user/hive/warehouse/bigdata.db/gd_ga_wa_wrysj/l/.hive-staging_hive_2019-03-16_08-10-08_139_4871120769486435512-104876/_tmp_-ext-10000
```

## Solution

1. Check whether the file is imported manually. If not, delete the file or directory.
2. If the file or directory cannot be deleted, change the file or directory permission to 770.

## 13.39 Failed to Load Data to Hive Tables

### Symptom

After creating a table, a user runs the **LOAD** command to import data to the table. However, the following problem occurs during the import:

```
.....
> LOAD DATA INPATH '/user/tester1/hive-data/data.txt' INTO TABLE employees_info;
Error: Error while compiling statement: FAILED: SemanticException Unable to load data to destination table.
Error: The file that you are trying to load does not match the file format of the destination table.
(state=42000,code=40000)
.....
```

### Cause Analysis

1. The storage format is not specified during table creation, and the default format RCFile is used.
2. However, the data to be imported is in TEXTFILE format.

### Solution

This problem is caused by an application defect. You can use a proper method based on site requirements only by ensuring that the storage format specified by the table is the same as the format of the data to be imported.

- Method 1:  
Specify the storage format when creating a table as a user who has the Hive table operation permission. For example:

```
CREATE TABLE IF NOT EXISTS employees_info(name STRING,age INT)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS
TEXTFILE;
```

Specify the format of the data to be imported as TEXTFILE.

- Method 2:  
Import RCFile data, but not TEXTFILE data.

## 13.40 Failed to Run the Application Developed Based on the Hive JDBC Code Case

### Symptom

After a user develops a service application by referring to the **jdbc-examples** sample project of the Hive component, the application fails to be executed. The application reports the following exception:

```
.....
2017-05-11 14:33:52.174 ERROR --- [          main] o.a.thrift.transport.TSaslTransport : SASL
negotiation failure
javax.security.sasl.SaslException: GSS initiate failed
at com.sun.security.sasl.gsskerb.GssKrb5Client.evaluateChallenge(Unknown Source)
at org.apache.thrift.transport.TSaslClientTransport.handleSaslStartMessage(TSaslClientTransport.java:94)
at org.apache.thrift.transport.TSaslTransport.open(TSaslTransport.java:271)
at org.apache.thrift.transport.TSaslClientTransport.open(TSaslClientTransport.java:37)
at org.apache.hadoop.hive.thrift.client.TUGIAssumingTransport$1.run(TUGIAssumingTransport.java:52)
at org.apache.hadoop.hive.thrift.client.TUGIAssumingTransport$1.run(TUGIAssumingTransport.java:49)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Unknown Source)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1711)
at org.apache.hadoop.hive.thrift.client.TUGIAssumingTransport.open(TUGIAssumingTransport.java:49)
at org.apache.hive.jdbc.HiveConnection.openTransport(HiveConnection.java:260)
at org.apache.hive.jdbc.HiveConnection.createClient(HiveConnection.java:213)
at org.apache.hive.jdbc.HiveConnection.<init>(HiveConnection.java:178)
at org.apache.hive.jdbc.HiveDriver.connect(HiveDriver.java:105)
at java.sql.DriverManager.getConnection(Unknown Source)
at java.sql.DriverManager.getConnection(Unknown Source)
at com.xxx.bigdata.hive.example.JDBCExample.main(JDBCExample.java:107)
Caused by: org.ietf.jgss.GSSException: No valid credentials provided (Mechanism level: Failed to find any
Kerberos tgt)
at sun.security.jgss.krb5.Krb5InitCredential.getInstance(Unknown Source)
at sun.security.jgss.krb5.Krb5MechFactory.getCredentialElement(Unknown Source)
at sun.security.jgss.krb5.Krb5MechFactory.getMechanismContext(Unknown Source)
at sun.security.jgss.GSSManagerImpl.getMechanismContext(Unknown Source)
at sun.security.jgss.GSSContextImpl.initSecContext(Unknown Source)
at sun.security.jgss.GSSContextImpl.initSecContext(Unknown Source)
... 17 common frames omitted
.....
```

### Cause Analysis

1. It is suspected that service interaction is performed before Kerberos authentication is complete.
2. Further analyze the logs. The log contains "com.xxx.bigdata.security.LoginUtil - Login success!!!!!!!!!!!!!!!" but not "org.apache.hadoop.security.UserGroupInformation : Login successful...".

Analyze the code. It is found that:

```
/* */ @InterfaceAudience.Public
/* */ @InterfaceStability.Evolving
/* */ public static synchronized void loginUserFromKeytab(String user, String path)
/* */ throws IOException
/* */ {
/* 958 */ if (!isSecurityEnabled()) {
/* 959 */ return;
```

```
/* */ }  
.....
```

3. Analyze **isSecurityEnabled()** and check whether **hadoop.security.authentication** is set to **kerberos** in the configuration.  
This Hive service application is not correctly configured. Therefore, the system determines that Kerberos authentication is not required.  
Analyze the **jdbc-examples** sample project of the Hive component. This problem does not occur in the sample project because the **core-site.xml** configuration file exists in the **classpath** directory of the project and **hadoop.security.authentication** is set to **kerberos** in the configuration file.

## Solution

Use any of the following methods to solve the problem:

- Method 1:  
Save the **core-site.xml** configuration file in the **classpath** directory by referring to the **jdbc-examples** sample project of the Hive component.

- Method 2:  
In the code, explicitly load the configuration file **core-site.xml**:

```
.....  
conf = new Configuration();  
String userdir = System.getProperty("user.dir") + File.separator + "conf" + File.separator;  
conf.addResource(new Path(userdir + "core-site.xml"));  
.....
```

- Method 3:  
In the code, set **hadoop.security.authentication** to **kerberos**:

```
.....  
CONF = new Configuration();  
CONF.set("hadoop.security.authentication", "kerberos");  
.....
```

## 13.41 HiveServer and HiveHCat Process Faults

### Issue

The HiveServer and WebHCat processes in the customer cluster are faulty.

### Symptom

The HiveServer and WebHCat processes on the Master2 node in the MRS cluster are faulty. After the restart, the processes are still faulty.

### Cause Analysis

On Manager, start the faulty HiveServer process. Log in to the node and search for the error information at the corresponding time point in the **hiveserver.out** log file. Error messages **error parsing conf mapred-site.xml** and **Premature end of file** are found. Restart WebHCat. The same error is reported because the **mapred-site.xml** file fails to be parsed.



## Procedure

1. Log in to the Master2 node as user **root**.
2. Run the **find / -name 'mapred-site.xml'** command to obtain the location of the **mapred-site.xml** file.
  - The path of HiveServer is **/opt/Bigdata/Cluster version/1\_13\_HiveServer/etc/mapred-site.xml**.
  - The path of WebHCat is **/opt/Bigdata/Cluster version/1\_13\_WebHCat/etc/mapred-site.xml**.
3. Check whether the **mapred-site.xml** file is normal. In this case, the configuration file is empty. As a result, the parsing fails.
4. Restore the **mapred-site.xml** file. Run the **scp** command to copy the configuration file in the corresponding directory on the Master1 node to the corresponding directory on the Master2 node to replace the original file.
5. Run the **chown omm:wheel mapred-site.xml** command to change the owner group and user.
6. On Manager, restart the faulty HiveServer and WebHCat processes.

## 13.42 Error Message "ConnectionLoss for hiveserver2" Is Displayed When MRS Hive Connects to ZooKeeper

### Symptom

Hive 1.2.1 of the MRS 1.8 cluster is successfully connected to the MRS cluster through the JDBC interface of Hive. However, an error is reported when Hive 2.3.2 of the MRS 1.9.0 cluster is connected to the MRS cluster through the JDBC interface of Hive for computing tasks.

An error message as shown in the following figure is displayed:

```
Caused by: org.apache.zookeeper KeeperException$ConnectionLossException: KeeperErrorCode = ConnectionLoss for /hiveserver2
```

### Cause Analysis

The authentication method of MRS 1.8 is the same as that of the open-source version. The Kerberos authentication file is directly transferred to the Hive URL for authentication. However, in MRS 1.9, this method cannot be used. You need to perform Kerberos authentication, obtain some other configuration information, and then combine the URL.

### Procedure

For details, see the authentication connection of **hive-examples** in the MRS 1.9 sample project. For details about the sample address, see the [Developer Guide](#).

### Summary and Suggestions

Develop code based on the official MRS 1.9 samples.

# 13.43 An Error is Reported When Hive Executes the insert into Statement

## Symptom

When a user uses MRS Hive to execute a SQL statement, the following error message is displayed.

Figure 13-1 Error reported when MRS Hive executes a SQL statement

```
0_762_995046968543258554-19104/-local-10004/HashTable-Stage-7/MapJoin_mpf1e121651-..._hashtable
2020-06-02 17:10:02   Uploaded 1 file to: file:/opt/bigdata/tmp/hive/localtmp/3c389d8-827f-4454-8Baa-c47e57127d9d/hive_2020-06-02-17-08-50_762_995046968543258554-19104/-local-10
HashTable-Stage-7/MapJoin_mpf1e121651-..._hashtable (104888 bytes)
2020-06-02 17:10:02   End of local task, Time Taken: 5.211 sec.
Error: org.apache.hive.service.cli.HiveSQLException: Error while processing statement: FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.ColumnStatsTask
at org.apache.hive.service.cli.operation.Operation.toSQLException(Operation.java:380)
at org.apache.hive.service.cli.operation.SQLOperation.runQuery(SQLOperation.java:268)
at org.apache.hive.service.cli.operation.SQLOperation.access$800(SQLOperation.java:93)
at org.apache.hive.service.cli.operation.SQLOperationsBackgroundWorker$1.run(SQLOperation.java:379)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1840)
at org.apache.hive.service.cli.operation.SQLOperationsBackgroundWorker.run(SQLOperation.java:393)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748) [state=96501,code=1]
```

## Cause Analysis

1. The HiveServer log shows the following message at the time when the error is reported.

Figure 13-2 HiveServer log

```
at org.apache.hadoop.hive.ql.Driver.run(Driver.java:1238)
at org.apache.hadoop.hive.ql.Driver.run(Driver.java:1233)
at org.apache.hive.service.cli.operation.SQLOperation.runQuery(SQLOperation.java:266)
at org.apache.hive.service.cli.operation.SQLOperation.access$800(SQLOperation.java:93)
at org.apache.hive.service.cli.operation.SQLOperationsBackgroundWorker$1.run(SQLOperation.java:379)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1840)
at org.apache.hive.service.cli.operation.SQLOperationsBackgroundWorker.run(SQLOperation.java:393)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
org.apache.hadoop.hive.ql.metadata.Hive.setPartitionColumnStatistics(Hive.java:378)
2020-06-02 16:11:03.771 | ERROR | HiveServer2-Background-Pool: Thread-2440344 | Failed to run column stats task | org.apache.hadoop.hive.ql.exec.ColumnStatsTask.execute(ColumnStatsTask.java:333)
org.apache.hadoop.hive.ql.metadata.HiveException: org.apache.thrift.transport.TTransportException
at org.apache.hadoop.hive.ql.exec.ColumnStatsTask.execute(ColumnStatsTask.java:420) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.exec.ColumnStatsTask.access$000(ColumnStatsTask.java:431) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.exec.Task.executeTask(Task.java:199) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.exec.TaskRunner.runSequential(TaskRunner.java:100) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.Driver.launchTask(Driver.java:218) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.Driver.run(Driver.java:1841) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.Driver.runInternal(Driver.java:1571) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.Driver.run(Driver.java:1238) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hive.service.cli.operation.SQLOperation.runQuery(SQLOperation.java:266) ~[hive-service-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hive.service.cli.operation.SQLOperationsBackgroundWorker$1.run(SQLOperation.java:379) ~[hive-service-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at java.security.AccessController.doPrivileged(Native Method) ~[?:1.8.0_232]
at javax.security.auth.Subject.doAs(Subject.java:422) ~[?:1.8.0_232]
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1840) ~[hadoop-common-3.0.3-mrs-1.9.0.jar:3]
at org.apache.hive.service.cli.operation.SQLOperationsBackgroundWorker.run(SQLOperation.java:393) ~[hive-service-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511) ~[?:1.8.0_232]
at java.util.concurrent.FutureTask.run(FutureTask.java:266) ~[?:1.8.0_232]
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149) ~[?:1.8.0_232]
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624) ~[?:1.8.0_232]
at java.lang.Thread.run(Thread.java:748) ~[?:1.8.0_232]
Caused by: org.apache.thrift.transport.TTransportException
at org.apache.thrift.transport.TIOStreamTransport.read(TIOStreamTransport.java:132) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:88) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TSaslTransport.readLength(TSaslTransport.java:376) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:450) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:435) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:37) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TSaslTransport.read(TSaslTransport.java:86) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.transport.TFramedTransport.readAll(TFramedTransport.java:62) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.protocol.TBinaryProtocol.readAll(TBinaryProtocol.java:449) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.protocol.TBinaryProtocol.readAll(TBinaryProtocol.java:218) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.protocol.TBinaryProtocol.readMessageBegin(TBinaryProtocol.java:219) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.TServiceClient.receiveBase(TServiceClient.java:77) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.metastore.api.ThriftHiveMetastoreClient.set_aggr_stats_for(ThriftHiveMetastoreClient.java:3573) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.metastore.api.HiveMetastoreClient.setAggrStatsFor(ThriftHiveMetastoreClient.java:3715) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.ql.metadata.SessionHiveMetastoreClient.setPartitionColumnStatistics(SessionHiveMetastoreClient.java:355) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at sun.reflect.GeneratedMethodAccessor19.invoke(Unknown Source) ~[?:?]
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43) ~[?:1.8.0_232]
at java.lang.reflect.Method.invoke(Method.java:498) ~[?:1.8.0_232]
at org.apache.hadoop.metastore.RetryingMetastoreClient.invoke(RetryingMetastoreClient.java:173) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at com.sun.proxy.$Proxy25.setPartitionColumnStatistics(Unknown Source) ~[?:?]
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43) ~[?:1.8.0_232]
at java.lang.reflect.Method.invoke(Method.java:498) ~[?:1.8.0_232]
at org.apache.hadoop.hive.metastore.HiveMetastoreClient$SynchronizedHandler.invoke(HiveMetastoreClient.java:2376) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at com.sun.proxy.$Proxy35.setPartitionColumnStatistics(Unknown Source) ~[?:?]
at org.apache.hadoop.hive.ql.metadata.Hive.setPartitionColumnStatistics(Hive.java:378) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
... 21 more
```

2. No important information is found in that log, but the **metadata** field is found in the stack. Therefore, the error may be related to MetaStore.

Figure 13-3 Metadata in the stack

```
2020-06-02 16:11:03.771 | ERROR | HiveServer2-Background-Pool: Thread-2440344 | Failed to run column stats task | org.apache.hadoop.hive.ql.exec.ColumnStatsTask.execute(ColumnStatsTask.java:333)
org.apache.hadoop.hive.ql.metadata.HiveException: org.apache.thrift.transport.TTransportException
at org.apache.hadoop.hive.ql.metadata.Hive.setPartitionColumnStatistics(Hive.java:378) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
```

3. The MetaStore log shows the following error information.

Figure 13-4 MetaStore log

```

[org.apache.hadoop.hive.metastore.RetryingHMSHandler.invokeInternal(RetryingHMSHandler.java:204)
Caused by: org.postgresql.util.PSQLException: ERROR: pool-12-thread-155 | Error occurred during processing of message. | org.apache.thrift.server.TThreadPoolServer$WorkerProcess.run(TThreadPoolServer.java:297)
org.datanucleus.exceptions.NucleusDataStoreException: Put request failed [UPDATE PARTITION_PARAMS SET PARAM_VALUE = ? WHERE PART_ID=? AND PARAM_KEY=?
at org.datanucleus.store.rdbms.scostore.JoinMapStore.put(JoinMapStore.java:318) ~[datanucleus-rdbms-4.1.19.jar:?]
at org.datanucleus.store.types.wrappers.BackedMap.put(Map.java:63) ~[datanucleus-core-4.1.17.jar:?]
at org.apache.hadoop.hive.common.StatsSetupConst.setColumnStats(StatsSetupConst.java:251) ~[hive-common-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.metastore.ObjectStore.updatePartitionColumnStatistics(ObjectStore.java:794) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at sun.reflect.GeneratedMethodAccessor99.invoke(Unknown Source) ~[?:?]
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43) ~[?:1.8.0_232]
at java.lang.reflect.Method.invoke(Method.java:498) ~[?:1.8.0_232]
at com.sun.proxy.$Proxy25.updatePartitionColumnStatistics(Unknown Source) ~[?:?]
at org.apache.hadoop.hive.metastore.RawStoreProxy.invoke(RawStoreProxy.java:101) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.updatePartitionColStats(HiveMetaStore.java:5138) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at sun.reflect.GeneratedMethodAccessor75.invoke(Unknown Source) ~[?:?]
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43) ~[?:1.8.0_232]
at java.lang.reflect.Method.invoke(Method.java:498) ~[?:1.8.0_232]
at org.apache.hadoop.hive.metastore.RetryingHMSHandler.invokeInternal(RetryingHMSHandler.java:148) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.metastore.RetryingHMSHandler.invoke(RetryingHMSHandler.java:107) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at com.sun.proxy.$Proxy33.set_aggr_stats_for(Unknown Source) ~[?:?]
at org.apache.hadoop.hive.metastore.api.ThriftHiveMetastoreProcessor$set_aggr_stats_for.getResult(ThriftHiveMetastore.java:13239) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.metastore.api.ThriftHiveMetastoreProcessor$set_aggr_stats_for.getResult(ThriftHiveMetastore.java:13233) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.ProcessFunction.process(ProcessFunction.java:39) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.TBaseProcessor.process(TBaseProcessor.java:29) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.thrift.HadoopThriftAuthBridge$Server$TUGIAssumingProcessor1.run(HadoopThriftAuthBridge.java:594) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.hadoop.hive.thrift.HadoopThriftAuthBridge$Server$TUGIAssumingProcessor1.run(HadoopThriftAuthBridge.java:589) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at java.security.AccessController.doPrivileged(Native Method) ~[?:1.8.0_232]
at javax.security.auth.Subject.doAs(Subject.java:422) ~[?:1.8.0_232]
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1840) ~[hadoop-common-2.8.3-mrs-1.9.0.jar:?]
at org.apache.hadoop.hive.thrift.HadoopThriftAuthBridge$Server$TUGIAssumingProcessor.process(HadoopThriftAuthBridge.java:589) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at org.apache.thrift.server.TThreadPoolServer$WorkerProcess.run(TThreadPoolServer.java:286) ~[hive-exec-2.3.3-mrs-1.9.0.jar:2.3.3-mrs-1.9.0]
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149) ~[?:1.8.0_232]
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624) ~[?:1.8.0_232]
at java.lang.Thread.run(Thread.java:748) ~[?:1.8.0_232]
Caused by: org.datanucleus.store.rdbms.exceptions.MappedDataStoreException: UPDATE PARTITION_PARAMS SET PARAM_VALUE = ? WHERE PART_ID=? AND PARAM_KEY=?
at org.datanucleus.store.rdbms.scostore.JoinMapStore.internalUpdate(JoinMapStore.java:1120) ~[datanucleus-rdbms-4.1.19.jar:?]
at org.datanucleus.store.rdbms.scostore.JoinMapStore.put(JoinMapStore.java:304) ~[datanucleus-rdbms-4.1.19.jar:?]
Caused by: org.postgresql.util.PSQLException: ERROR: value too long for type character varying(4000)
at org.postgresql.core.v3.QueryExecutorImpl.receiveErrorResponse(QueryExecutorImpl.java:199) ~[postgresql-9.4.12000.jar:9.4.12000]
at org.postgresql.core.v3.QueryExecutorImpl.processResults(QueryExecutorImpl.java:1928) ~[postgresql-9.4.12000.jar:9.4.12000]
at org.postgresql.core.v3.QueryExecutorImpl.execute(QueryExecutorImpl.java:348) ~[postgresql-9.4.12000.jar:9.4.12000]
at org.postgresql.jdbc2.AbstractJdbc2Statement.execute(AbstractJdbc2Statement.java:545) ~[postgresql-9.4.12000.jar:9.4.12000]
at org.postgresql.jdbc2.AbstractJdbc2Statement.executeUpdate(AbstractJdbc2Statement.java:419) ~[postgresql-9.4.12000.jar:9.4.12000]
at org.postgresql.jdbc2.AbstractJdbc2Statement.executeUpdate(AbstractJdbc2Statement.java:365) ~[postgresql-9.4.12000.jar:9.4.12000]
at org.datanucleus.store.rdbms.ParamLoggingPreparedStatement.executeUpdate(ParamLoggingPreparedStatement.java:393) ~[datanucleus-rdbms-4.1.19.jar:?]
at org.datanucleus.store.rdbms.SQLController.executeUpdate(SQLController.java:431) ~[datanucleus-rdbms-4.1.19.jar:?]
at org.datanucleus.store.rdbms.scostore.JoinMapStore.internalUpdate(JoinMapStore.java:1010) ~[datanucleus-rdbms-4.1.19.jar:?]
at org.datanucleus.store.rdbms.scostore.JoinMapStore.put(JoinMapStore.java:304) ~[datanucleus-rdbms-4.1.19.jar:?]
... 30 more
2020-06-02 16:19:28.125 | INFO | pool-12-thread-155 | 156: Cleaning up thread local RawStore... | org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.logInfo(HiveMetaStore.java:885)

```

The error context indicates that an error occurs during SQL statement execution, and the following information is displayed in the error message:

Caused by: org.postgresql.util.PSQLException: ERROR: value too long for type character varying(4000)

The SQL statement fails because the length of all columns exceeds 4000 bytes. The restriction needs to be modified.

## Procedure

**Step 1** Log in to any master node in the cluster as user **root** and run the **su - omm** command to switch to user **omm**.

**Step 2** Run the following command to log in to the database. There can be security risks if a command contains the authentication password. You are advised to disable the command recording function (history) before running the command.

```
gsql -p 20051 -d hivemeta -U username -W password
```

**Step 3** Run the following command to modify the restriction:

```
alter table PARTITION_PARAMS alter column PARAM_VALUE type  
varchar(6000);
```

```
----End
```

## 13.44 Timeout Reported When Adding the Hive Table Field

### Issue

An error message is reported when adding the Hive table fields.

## Symptom

Hive executes **ALTER TABLE table\_name ADD COLUMNS(column\_name string) CASCADE** on tables that contain more than 10,000 partitions. The error information is as follows:

```
Timeout when executing method: alter_table_with_environment_context; 600525ms exceeds 600000ms
```

## Cause Analysis

1. The MetaStore client connection times out. The default timeout interval for the connection between the MetaStore client and server is 600 seconds. On Manager, increase the value of **hive.metastore.client.socket.timeout** to **3600s**.
2. Another error is reported:  
Error: org.apache.hive.service.cli.HiveSQLException: Error while processing statement: FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask. Unable to alter table. java.net.SocketTimeoutException: Read timed out  
JDBC connection timeout interval of the MetaStore metadata. The default value is 60 ms.
3. Increase the value of **socketTimeout** in **javax.jdo.option.ConnectionURL** to **60000**. The initial error is still reported.  
Timeout when executing method: alter\_table\_with\_environment\_context;3600556ms exceeds 3600000ms
4. Increase the values of parameters such as **hive.metastore.batch.retrieve.max**, **hive.metastore.batch.retrieve.table.partition.max**, and **dbservice.database.max.connections**. The problem persists.
5. It is suspected that the problem is caused by the GaussDB because adding a field will traverse each partition to execute **getPartitionColumnStatistics** and **alterPartition**.
6. Run the **gsql -p 20051 -U omm -W password -d hivemeta** command as user **omm** to log in to the Hive metabase. There can be security risks if a command contains the authentication password. You are advised to disable the command recording function (history) before running the command.
7. Run **select \* from pg\_locks**. No lock wait is found.
8. Run **select \* from pg\_stat\_activity**; It is found that the process execution takes a long time.

```
SELECT 'org.apache.hadoop.hive.metastore.model.MPartitionColumnStatistics'AS  
NUCLEUS_TYPE,A0.AVG_COL_LEN,A0."COLUMN_NAME",A0.COLUMN_TYPE,A0.DB_NAME,A0.BIG_DECI  
MAL_HIGH_VALUE,A0.BIG_DECIMAL_LOW_VALUE,A0.DOUBLE_HIGH_VALUE,A0.DOUBLE_LOW_VALUE,  
A0.LAST_ANALYZED,A0.LONG_HIGH_VALUE,A0.LONG_LOW_VALUE,A0.MAX_COL_LEN,A0.NUM_DISTIN  
CTS,A0.NUM_FALSES,A0.NUM_NULLS,A0.NUM_TRUES,A0.PARTITION_NAME,A0."TABLE_NAME",A0.CS_  
ID,A0.PARTITION_NAMEAS NUCORDER0 FROM PART_COL_STATS A0 WHERE A0."TABLE_NAME" =  
'$1' ANDA0.DB_NAME = '$2' AND A0.PARTITION_NAME = '$3' AND((((A0."COLUMN_NAME" = '$4')  
OR (A0."COLUMN_NAME" = '$5')) OR (A0."COLUMN_NAME" = '$6')) OR (A0."COLUMN_NAME"  
= '$7')) OR (A0."COLUMN_NAME" = '$8')) OR (A0."COLUMN_NAME" = '$9')) ORDER BY NUCORDER0;
```

```

JOB_ID	USER_NAME	PID	QUERY_ID	USERNAME	APPLICATION_NAME	CLIENT_ADDR	CLIENT_PORT	BACKING_START
FACT_START | STATE_CHANGE | QUERY_START | STATE | WAITING | STATE | CLIENT_HOSTNAME | CLIENT_PORT | BACKING_START |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
QUERY
-----|-----|-----|-----|-----|-----|-----|-----|-----|
28463 | POSTGRES | 10662 | 18 | omm | JobScheduler | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 32479 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 18402 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 27410 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 782 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1031 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1031 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1497 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1562 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 4421 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1521 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1703 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
1734A | SQOOP | 1553 | 1734A | SQOOP | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1654 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
1734A | SQOOP | 1678 | 1734A | SQOOP | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 20786 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 20880 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|
18396 | HIVEMETA | 1048 | 10393 | HIVE | | | | | |
-----|-----|-----|-----|-----|-----|-----|-----|-----|

```

9. Run the `gs_guc reload -c log_min_duration_statement=100 -D /srv/BigData/dbdata_service/data/` command to start SQL recording. It is found that the statement execution duration in **8** is 700 ms, and more than 10,000 commands are executed because there are more than 10,000 partitions.

10. Add **explain (analyze, verbose, timing, costs, buffers)** before the SQL statement to analyze the execution plan. It is found that the entire table needs to be scanned during execution.

```

hive> explain (analyze, verbose, timing, costs, buffers) SELECT * FROM sqs_active FROM sqs_active_model WHERE col_cd = '003240' FETCH NEXT ROW ONLY;
-----|-----|-----|-----|-----|-----|-----|-----|-----|
QUERY PLAN
-----|-----|-----|-----|-----|-----|-----|-----|-----|
Limit (cost=0.00..208.22 rows=1 width=218) (actual time=36.084..36.085 rows=1 loops=1)
-> Sort (cost=0.00..208.22 rows=1 width=218) (actual time=36.084..36.085 rows=1 loops=1)
  Output: (cost=0.00..208.22 rows=1 width=218) (actual time=36.084..36.085 rows=1 loops=1)
  Filter: (col_cd = '003240') (cost=0.00..208.22 rows=1 width=218) (actual time=36.084..36.085 rows=1 loops=1)
  Buffers: shared hit=8720
-> Seq Scan on sqs_active FROM sqs_active_model WHERE col_cd = '003240' (cost=0.00..208.22 rows=1 width=218) (actual time=36.084..36.085 rows=1 loops=1)
  Filter: (col_cd = '003240') (cost=0.00..208.22 rows=1 width=218) (actual time=36.084..36.085 rows=1 loops=1)
  Buffers: shared hit=8720
Total runtime: 36.143 ms
(9 rows)

```

11. Check the index. It is found that the index does not meet the leftmost match rule.

```

HIVEMETA=# \d+ PART_COL_STATS
-----|-----|-----|-----|-----|-----|-----|-----|-----|
Table "PUBLIC.PART_COL_STATS"
-----|-----|-----|-----|-----|-----|-----|-----|-----|
Column | Type | Modifiers | Storage | Stats target | Description
-----|-----|-----|-----|-----|-----|-----|-----|-----|
CS_ID | BIGINT | not null | plain | | 
CAT_NAME | CHARACTER VARYING(256) | default NULL::CHARACTER VARYING | extended | | 
DB_NAME | CHARACTER VARYING(128) | default NULL::CHARACTER VARYING | extended | | 
TABLE_NAME | CHARACTER VARYING(256) | default NULL::CHARACTER VARYING | extended | | 
PARTITION_NAME | CHARACTER VARYING(767) | default NULL::CHARACTER VARYING | extended | | 
COLUMN_NAME | CHARACTER VARYING(767) | default NULL::CHARACTER VARYING | extended | | 
COLUMN_TYPE | CHARACTER VARYING(128) | default NULL::CHARACTER VARYING | extended | | 
PART_ID | BIGINT | not null | plain | | 
LONG_LOW_VALUE | BIGINT | | plain | | 
LONG_HIGH_VALUE | BIGINT | | plain | | 
DOUBLE_LOW_VALUE | DOUBLE PRECISION | | plain | | 
DOUBLE_HIGH_VALUE | DOUBLE PRECISION | | plain | | 
BIG_DECIMAL_LOW_VALUE | CHARACTER VARYING(4000) | default NULL::CHARACTER VARYING | extended | | 
BIG_DECIMAL_HIGH_VALUE | CHARACTER VARYING(4000) | default NULL::CHARACTER VARYING | extended | | 
NUM_NULLS | BIGINT | not null | plain | | 
NUM_DISTINCTS | BIGINT | | plain | | 
BIT_VECTOR | BYTEA | | extended | | 
AVG_COL_LEN | DOUBLE PRECISION | | plain | | 
MAX_COL_LEN | BIGINT | | plain | | 
NUM_TRUES | BIGINT | | plain | | 
NUM_FALSES | BIGINT | | plain | | 
LAST_ANALYZED | BIGINT | not null | plain | | 
Indexes:
"PART_COL_STATS_pkey" PRIMARY KEY, BTREE (CS_ID)
"PART_COL_STATS_M49" BTREE (PART_ID)
"PCS_STATS_IDX" BTREE (CAT_NAME, DB_NAME, TABLE_NAME, COLUMN_NAME, PARTITION NAME)
Foreign-key constraints:
"PART_COL_STATS_fkey" FOREIGN KEY (PART_ID) REFERENCES PARTITIONS(PART_ID) DEFERRABLE
Has OIDs: no

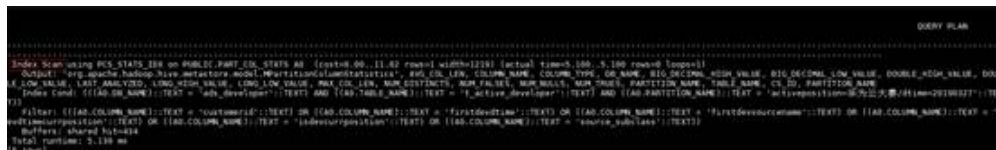
```

## Procedure

1. Rebuild an index.  
`su - omm`

```
gsql -p 20051 -U omm -W password -d hivemeta
DROP INDEX PCS_STATS_IDX;
CREATE INDEX PCS_STATS_IDX ON PART_COL_STATS(DB_NAME,
TABLE_NAME, COLUMN_NAME, PARTITION_NAME);
CREATE INDEX SDS_N50 ON SDS(CD_ID);
```

2. Check the execution plan again. It is found that the statement can be indexed and executed within 5 ms (the original execution time is 700 ms). Add fields to the Hive table again. The fields can be added successfully.



## 13.45 Failed to Restart Hive

### Issue

After the Hive service configuration is modified, the configuration fails to be saved. The configuration status of the Hive service on Manager is **Failed**.

### Symptom

User A opens the Hive configuration file in the background of the MRS node and does not close the file. User B modifies the Hive configuration item in **Service Management** on the MRS Manager page, saves the configuration, and restarts the Hive service. However, the configuration fails to be saved and the Hive service fails to be started.

### Cause Analysis

When user B modifies the configuration on the MRS Manager page, the configuration file is opened by user A in the background of an MRS node. As a result, the configuration file cannot be replaced and the Hive service fails to be started.

### Procedure

- Step 1** Manually close the Hive configuration file opened in the background of the cluster node.
- Step 2** Modify the Hive configuration on MRS Manager and save the configuration.
- Step 3** Restart the Hive service.

----End

## 13.46 Failed to Delete a Table Due to Excessive Hive Partitions

### Symptom

Partitioning a Hive table by two columns may eventually generate over 20,000 partition files. As a result, the user fails to execute the **truncate table** `{TableName}` or **drop table** `{TableName}` statement to delete table data.

### Cause Analysis

The file deletion operations are executed by a single thread serially. If the Hive partitioned tables have too many partition files, a large amount of metadata is stored in the metadata database. It takes a long time to delete metadata when a statement is executed to delete table data. As a result, the deletion cannot be complete within the specified timeout period, and the operation fails.

#### NOTE

You can log in to FusionInsight Manager and choose **Cluster > Services > Hive**. On the Hive page, choose **Configuration > All Configurations**, choose **ServerInit** under **MetaStore(Role)** in the navigation tree, and view the **hive.metastore.client.socket.timeout** parameter value in the right pane. This value is the timeout period. You can view the default value in the **Description** column.

### Procedure

- Step 1** (Optional, perform this step for an internal table) Use **alter table** `{TableName}` **set TBLPROPERTIES('EXTERNAL'='true')** to convert it into an external table. In this way, only its metadata but not data stored on the HDFS is deleted, saving the table deletion time.
- Step 2** (Optional, perform this step to use the same table name) Run the **show create table** `{TableName}` command to export the table structure, and then run the **ALTER TABLE** `{TableName}` **RENAME TO** `{new_table_name}`; command to rename the table. In this way, you can create a table that is the same as the original one.
- Step 3** Run the **hdfs dfs -rm -r -f** `{hdfs_path}` command to delete table data from HDFS.
- Step 4** Use **alter table** `{Table_Name}` **drop partition** (`{PartitionName}<'XXXX', {PartitionName}>'XXXX'`); in Hive to delete partitions and reduce the number of files. The deletion conditions can be flexibly configured.
- Step 5** When the number of rest partitions is smaller than 1,000, run the **drop table** `{TableName}` command to delete the table.

----End





## Procedure

- Step 1** Configure the Yarn queue permission for the current user.
- On Manager, bind the current user to a role with the corresponding permission, for example, **default**. For details, see [Modifying User Information](#).
  - If Ranger authentication is enabled for Hive, you can configure the permission for the current user to submit jobs in the specified Yarn queue by referring to [Adding a Ranger Access Permission Policy for Yarn](#).
- Step 2** After the permission configuration takes effect, run the command again.
- End

## 13.49 Releasing Disk Space After Dropping a Table in Hive

### Issue

After a user runs the **drop** command on the Hive CLI to drop a table and then uses the **hdfs dfsadmin -report** command to check the disk space, the command output shows that the table is not deleted.

### Cause Analysis

The **drop** command executed on the Hive CLI deletes only the table structure of the external table, but not the table data stored in HDFS.

### Procedure

- Step 1** Log in to the node where the client is installed as user **root** and authenticate the component user.
- ```
cd Client installation directory  
source bigdata_env  
init Component service user (Skip this step for clusters with Kerberos authentication disabled.)
```
- Step 2** Run the following command to delete the table stored in HDFS:
- ```
hadoop fs -rm hdfs://hacluster/Path of the table
```
- End

## 13.50 Abnormal Hive Query Due to Damaged Data in the JSON Table

### Issue

A user fails to query JSON data on the Hive client because the JSON table contains damaged data.

1. On the Hive client, the user executes the default open-source serialized JSON statement to create a table.

```
create external table if not exists test (  
name string  
)  
row format serde 'org.apache.hive.hcatalog.data.JsonSerDe'  
stored as textfile  
location 'hdfs://hacluster/user/hive/warehouse/database/table/jsondata';
```

2. Failed to query the table.

```
INFO : State: Compiling.  
INFO : Compiling command(queryId=omn_20221014092819_d396e069-8010-415c-a6e9-a7fa1191fb95): select * from test; Current sessionId=0139cc48-e14f-474e-9408-3920197d156e  
INFO : hive.compile.auto.avoid.cbo=true  
INFO : Concurrency mode is disabled, not creating a lock manager  
INFO : Current sql is not contains insert syntax, not need record dest table flag  
INFO : Semantic Analysis Completed (retrial = false)  
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name=test.name, type:string, comment:null)], properties:null)  
INFO : Completed compiling command(queryId=omn_20221014092819_d396e069-8010-415c-a6e9-a7fa1191fb95): Time taken: 0.037 seconds  
INFO : Concurrency mode is disabled, not creating a lock manager  
INFO : State: Executing.  
INFO : Executing command(queryId=omn_20221014092819_d396e069-8010-415c-a6e9-a7fa1191fb95): select * from test; Current sessionId=0139cc48-e14f-474e-9408-3920197d156e  
INFO : Completed executing command(queryId=omn_20221014092819_d396e069-8010-415c-a6e9-a7fa1191fb95): Time taken: 0.0 seconds  
INFO : OK  
INFO : Concurrency mode is disabled, not creating a lock manager  
Error: java.io.IOException: org.apache.hadoop.hive.serde2.SerdeException: org.codehaus.jackson.JsonParseException: Unexpected character ('1' (code 49)): was expecting a colon to separate field name and value  
at [Source: java.io.ByteArrayInputStream@1e9b799d; line: 1, column: 13] (state=,code=0)
```

### NOTE

This section applies to versions later than MRS 1.9.2.

### Cause Analysis

When the default open-source serialized JSON statement is used to create a Hive table, damaged JSON data in the table cannot be filtered out. As a result, the query is abnormal.

### Procedure

- Step 1** Download [json-serde-1.3.8-jar-with-dependencies.jar](#) to the local host and upload the JAR file to all HiveServer nodes.
- Step 2** Log in to the nodes where the JAR package is uploaded in [Step 1](#) and the Hive and HDFS clients are installed as the client installation user.
- Step 3** Run the following commands to authenticate the user:

```
cd Client installation directory  
source bigdata_env
```

**kinit** *User who has the Hive administrator rights* (Skip this operation for clusters with Kerberos authentication disabled)

**Step 4** Run the following command to create an HDFS directory for storing JAR files:

```
hdfs dfs -mkdir HDFS directory for storing JAR files
```

**Step 5** Run the following command to upload the JAR files in **Step 2** to HDFS:

```
hdfs dfs -put JAR file storage path HDFS directory for storing JAR files
```

**Step 6** Run the following commands to enable Hive to load a specified JAR file when executing a CLI task:

```
beeline
```

```
set role admin; (Skip this step for clusters with Kerberos authentication disabled)
```

```
add jar HDFS directory for storing JAR files
```

**Step 7** Run the following commands to create the table again:

```
create external table if not exists test (
```

```
name string
```

```
)
```

```
row format serde 'org.openx.data.jsonserde.JsonSerDe'
```

```
stored as textfile
```

```
location 'hdfs://hacluster/user/hive/warehouse/database/table/jsondata';
```

**Step 8** Run the following command to modify the table property to ignore damaged JSON data:

```
ALTER TABLE test SET SERDEPROPERTIES( "ignore.malformed.json" = "true");
```

After this property is modified, **NULL** is displayed by default if the queried table contains damaged JSON data.

**Step 9** Run the following command to query table data:

```
select * from test;
```

If the data query is successful, the damaged JSON data is ignored. The damaged JSON data is displayed as **NULL**.

```
INFO : Concurrency mode is disabled, not creating a lock manager
INFO : Current sql is not contains insert syntax, not need record dest table flag
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:[FieldSchema(name=test.name, type:string, comment:null)], properties:null)
INFO : Completed compiling command(queryId=ommm_20221014093100_2ad45753-8e44-43f6-8931-7c8d0c95f420); Time taken: 0.053 seconds
INFO : Concurrency mode is disabled, not creating a lock manager
INFO : State: Executing.
INFO : Executing command(queryId=ommm_20221014093100_2ad45753-8e44-43f6-8931-7c8d0c95f420): select * from test; Current sessionId=03156e
INFO : Completed executing command(queryId=ommm_20221014093100_2ad45753-8e44-43f6-8931-7c8d0c95f420); Time taken: 0.0 seconds
INFO : OK
INFO : Concurrency mode is disabled, not creating a lock manager
+-----+
|      name      |
+-----+
| 123            |
| NULL           |
+-----+
```

----End

## 13.51 Connection Timed Out During SQL Statement Execution on the Hive Client

### Symptom

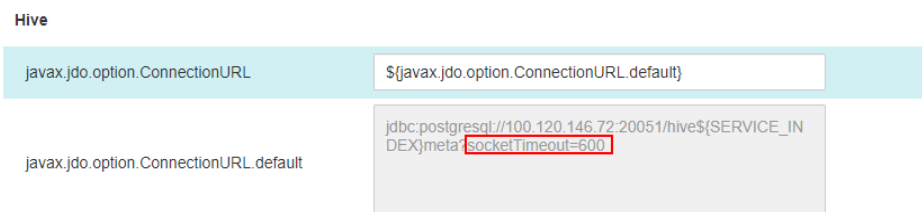
The Hive client fails to execute SQL statements, and error message "Timed out waiting for a free available connection" is displayed.

### Cause Analysis

There are a large number of DBService connections, and obtaining the connections times out.

### Procedure

- Step 1** Check whether the client uses the Spark SQL client to execute SQL statements.
- If yes, check the timeout parameter in the URL, change the value to **600**, and go to [Step 7](#).
  - If no, go to [Step 2](#).
- Step 2** Log in to Manager, choose **Cluster > Services > Hive > Configurations > All Configurations**, search for **javax.jdo.option.ConnectionURL**, and check whether the value of the timeout parameter is less than **600**.



#### NOTE

Hive, HiveServer, MetaStore, and WebHCat all have this parameter. Ensure that their parameter values are the same.

- If yes, go to [Step 3](#).
  - If no, go to [Step 7](#).
- Step 3** Check whether the value of **javax.jdo.option.ConnectionURL** is **\$(javax.jdo.option.ConnectionURL.default)**.
- If yes, go to [Step 4](#).
  - If no, change the timeout parameter in the URL to **600**, click **Save**, and go to [Step 7](#).
- Step 4** Click **Instance**, select any HiveServer instance, and log in to the instance node as user **root**.
- Step 5** Open the **`\${BIGDATA\_HOME}/FusionInsight\_Current/\*HiveServer/etc/hivemetastore-site.xml** configuration file, find the **javax.jdo.option.ConnectionURL** parameter, and copy its value.

```
<property>
<name>javax.jdo.option.ConnectionURL</name>
<value>jdbc:postgresql://100.120.146.72:20051/hivemeta?socketTimeout=60</value>
</property>
</property>
```

**Step 6** Log in to Manager, choose **Cluster > Services > Hive > Configurations > All Configurations**, search for **javax.jdo.option.ConnectionURL**, change its value to the URL copied in **Step 5**, change the timeout parameter to **600**, and click **Save**.

 **NOTE**

Hive, HiveServer, MetaStore, and WebHCat all have this parameter. Ensure that their parameter values are the same.

**Step 7** Choose **Cluster > Services > Hive > Configurations > All Configurations**, search for **maxConnectionsPerPartition**, and check whether its value is less than **100**.

- If yes, change the value to **100**, click **Save**, and go to **Step 8**.
- If no, go to **Step 8**.

**Step 8** If parameters are modified in the preceding steps, choose **Cluster > Services > Hive > Dashboard**, click **More**, and select **Service Rolling Restart**. If no parameters are modified, skip this step.

----End

## 13.52 WebHCat Failed to Start Due to Abnormal Health Status

### Issue

The WebHCat instance fails to be started.

### Symptom

On Manager, the health status of the WebHCat instance is **Faulty**, and alarm **ALM-12007 Process Fault** is generated for the WebHCat instance of the Hive service. An error is reported when the Hive service is restarted.

Error messages "Service not found in Kerberos database" and "Address already in use" are contained in the **/var/log/Bigdata/hive/webhcat/webhcat.log** file of the WebHCat instance.

### Procedure

**Step 1** Log in to each node where the WebHCat instance resides and check whether the mapping between IP addresses and hostnames in the **/etc/hosts** file is correct. The WebHCat configurations in the **/etc/hostname** and **/etc/HOSTNAME** files must be the same as those in the **/etc/hosts** file. If they are different, manually modify them.

 **NOTE**

To view the mapping between the IP addresses and hostnames of the WebHCat instance, log in to FusionInsight Manager and choose **Cluster > Services > Hive > Instance**.

**Step 2** Log in to any node where the WebHCat instance resides and run the following command to switch to user **omm**:

```
su - omm
```

**Step 3** Run the following command to check whether the WebHCat process exists:

```
ps -ef|grep webhcat|grep -v grep
```

If it does, run the following command to kill it:

```
kill -9 ${webhcat_pid}
```

**Step 4** Log in to FusionInsight Manager and choose **Cluster > Services > Hive** . On the page that is displayed, click the **Instance** tab. The select all WebHCat instances, click **More**, and select **Restart Instance**. Wait until WebHCat is restarted successfully.

----End

## 13.53 WebHCat Failed to Start Because the mapred-default.xml File Cannot Be Parsed

### Issue

The Hive service of MRS is faulty. After the Hive service is restarted, the HiveServer and WebHCat processes on the Master2 node fail to start, but the processes on the Master1 node are normal.

### Cause Analysis

Log in to the Master2 node and check the `/var/log/Bigdata/hive/hiveserver/hive.log` file. It is found that HiveServer keeps loading `/opt/Bigdata/*/*_HiveServer/etc/hive-site.xml`. Check the `/var/log/Bigdata/hive/hiveserver/hiveserver.out` log generated when HiveServer exits. It is found that an exception occurs when the `mapred-default.xml` file is parsed.

### Procedure

**Step 1** Log in to the Master2 node and run the following command to query the path of `mapred-default.xml`:

```
find /opt/ -name 'mapred-default.xml'
```

The configuration file is in the `/opt/Bigdata/*/*_WebHCat/etc/` directory but is empty.

**Step 2** Log in to the Master1 node, copy the `/opt/Bigdata/*/*_WebHCat/etc/mapred-default.xml` file to the Master2 node, and change the owner group of the file to `omm:wheel`.

**Step 3** Log in to Manager and restart the abnormal HiveServer and WebHCat instances.

----End

## 13.54 An SQL Error Is Reported When the Number of MetaStore Dynamic Partitions Exceeds the Threshold

### Symptom

When the SparkSQL or HiveSQL command is executed, the following error message is displayed:

```
Number of dynamic partitions created is 2001, which is more than 2000. To solve this try to set hive.exec.max.dynamic.partitions to at least 2001
```

### Cause Analysis

By default, Hive limits the maximum number of dynamic partitions, which is specified by the **hive.exec.max.dynamic.partitions** parameter. The default value is **1000**. If the threshold is exceeded, Hive will not create new dynamic partitions.

### Procedure

- Adjust upper-layer services to ensure that the number of dynamic partitions is within the value range of **hive.exec.max.dynamic.partitions**.
- Run the **set hive.exec.max.dynamic.partitions = XXX;** command to increase the value of **hive.exec.max.dynamic.partitions**.

The **spark.hadoop.hive.exec.max.dynamic.partitions** parameter needs to be set in SparkSQL.

# 14 Using Hue

## 14.1 An Unknown Job Is Running on the Hue Page

### Issue

An unknown job is running on the Hue page.

### Symptom

After the MRS cluster is created, a job that is not submitted by the user is running on Hue.

Job ID	Job Name	Job Type	Status	Progress	Priority	Queue	Yarn	Start Time	End Time	
132242208965_2008	select count(*) from tab_lockkeeper(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	17s	07/26/18 11:22:13	
132242208965_2007	select count(*) from tab_lockkeeper(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	20s	07/26/18 11:22:34	
132242208965_2004	select count(*) from tab_lockkeeper(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	20s	07/26/18 11:22:47	
132242208965_2005	select count(*) from tab_lockkeeper(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	19s	07/26/18 08:25:18	
132242208965_2004	select count(*) from tab_lockkeeper(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	24s	07/26/18 08:38:36	
132242208965_2003	select count(*) from tab_lockkeeper(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	23s	07/26/18 08:46:25	
132242208965_2002	select count(*) from TAB_LOCKKEEPER(Stage 1)	MAPREDUCE	SUCCESS	100%	100%	default	无	19s	07/26/18 20:01:00	
132242208965_2001	Spark JDBCServer 192.168.1.163	SPARK	Running	100%	100%	default	无	22s 12m 18s	07/26/18 17:14:41	
132242208965_2001	Spark JDBCServer 192.168.1.163	SPARK	SUCCESS	100%	100%	default	无	30m 37s	07/26/18 16:33:23	
132242208965_2001	Spark JDBCServer 192.168.1.163	SPARK	SUCCESS	100%	100%	default	无	40 45m 31s	07/26/18 09:43:32	

### Cause Analysis

This job is a default permanent job generated when the system connects to JDBC after Spark is started.

### Procedure

This is not a problem. No handling is required.



## 14.2 HQL Fails to Be Executed on Hue Using Internet Explorer

### Symptom

Using Internet Explorer to access Hive Editor and execute all HQL statements on Hue fails and the system prompts "There was an error with your query".

### Cause Analysis

Internet Explorer has functional problems and cannot process AJAX POST requests containing form data in 307 redirection. Use a compatible browser.

### Solution

Google Chrome 21 or later is recommended.

## 14.3 Failed to Access the Hue Web UI

### Issue

An error page is displayed when the Hue web UI is accessed.

### Symptom

The following error information is displayed for access to the Hue web UI:

```
503 Service Unavailable
The server is temporarily unable to service your requester due to maintenance downtime or capacity
problems.Please try again later.
```

### Cause Analysis

The floating IP address of Hue is faulty.

### Procedure

**Step 1** Log in to Manager, choose **Cluster > Services > Hue > Configurations**, search for **HUE\_FLOAT\_IP**, and view the floating IP address configured for Hue.

**Step 2** Connect to the active Hue instance node and check whether the active instance is bound to the NIC of the floating IP address.

```
ifconfig | grep -A2 HUE | grep <Floating IP address>
```

- If yes, the command output is not empty and the active instance is bound to the floating IP address. Then, go to [Step 4](#).
- If no, the command output is empty and the active instance is not bound to the floating IP address. Then, go to [Step 3](#).

**Step 3** Connect to the active Hue instance node and bind it to the NIC of the floating IP address. You can obtain the NIC, floating IP address, and subnet mask of the node from the NIC information corresponding to the local IP address in the **ifconfig** command output.

```
ifconfig <NIC of the node>:HUE <Floating IP address> netmask <Subnet mask>
```

**Step 4** Connect to the active OMS node, try to connect to the floating IP address, and check whether the current host name is the name of the active Hue instance node.

```
su - omm
```

```
ssh <Floating IP address>
```

- If yes, the floating IP address is normal. If the Hue web UI still cannot be accessed, contact technical support.
- If no, the floating IP address is occupied. Go to [Step 5](#).

**Step 5** Log in to all other nodes except the active Hue instance in the cluster and run the following command to check whether they occupy the floating IP address of Hue:

```
ifconfig | grep -B2 <Floating IP address>
```

- If yes, run the **ifconfig** <NIC name > **down** command to unbind the NIC and then go to [Step 6](#).
- If no, contact technical support.

**Step 6** Repeat [Step 4](#) to [Step 5](#) until you can connect to the active Hue instance using the floating IP address and access the Hue web UI.

----End

## 14.4 HBase Tables Cannot Be Loaded on the Hue Web UI

### Issue

After Hive data is imported to HBase on the Hue web UI, an error message is displayed, indicating that the HBase table cannot be detected.

### Symptom

In the Kerberos cluster, the IAM sub-account does not have sufficient permissions to load HBase tables.

### Cause Analysis

The IAM sub-account does not have sufficient permissions.

### Procedure

MRS Manager:

**Step 1** Log in to MRS Manager.

**Step 2** Choose **System > Manage User**.

**Step 3** Locate the row that contains the target user, and click **Modify**.

**Step 4** Add the user to the **supergroup** group.

**Step 5** Click **OK**.

----End

FusionInsight Manager:

**Step 1** Log in to FusionInsight Manager.

**Step 2** Choose **System > Permission > User**.

**Step 3** Locate the row that contains the target user, and click **Modify**.

**Step 4** Add the user to the **supergroup** group.

**Step 5** Click **OK**.

----End

## Summary and Suggestions

If Kerberos authentication is enabled for a cluster, "No data available" is displayed on the page. In this case, check the permission first.

# 15 Using Impala

## 15.1 Failed to Connect to impala-shell

### Issue

A user fails to connect to impala-shell.

### Symptom

After a user modifies the configuration of any component on the component management page and restarts the service, the connection to impala-shell fails, and the error message "no such file/directory" is displayed.

```
[root@node-master1emdj etc]# pwd
/opt/Bigdata/MRS_2.1.0/1_7_KuduMaster/etc
[root@node-master1emdj etc]# impala-shell -i 192.168.0.73
shell-init: error retrieving current directory: getcwd: cannot access parent directories: No such file or directory
chdir: error retrieving current directory: getcwd: cannot access parent directories: No such file or directory
Traceback (most recent call last):
  File "/opt/client/Impala/impala/shell/impala_shell.py", line 38, in <module>
    from impala_client import (ImpalaClient, DisconnectedException, QueryStateException,
  File "/opt/client/Impala/impala/shell/lib/impala_client.py", line 20, in <module>
    import sasl
  File "build/bdist.linux-x86_64/egg/sasl/_init_.py", line 1, in <module>
    File "build/bdist.linux-x86_64/egg/sasl/saslwrapper.py", line 7, in <module>
  File "build/bdist.linux-x86_64/egg/_saslwrapper.py", line 7, in <module>
  File "build/bdist.linux-x86_64/egg/_saslwrapper.py", line 3, in __bootstrap__
  File "/usr/lib/python2.7/site-packages/setuptools-0.6c11-py2.7.egg/pkg_resources.py", line 2594, in <module>
    for comparator, version in req.specs:
  File "/usr/lib/python2.7/site-packages/setuptools-0.6c11-py2.7.egg/pkg_resources.py", line 425, in __init__
    File "/usr/lib/python2.7/site-packages/setuptools-0.6c11-py2.7.egg/pkg_resources.py", line 440, in add_entry
    `req`. But, if there is an active distribution for the project and it
  File "/usr/lib/python2.7/site-packages/setuptools-0.6c11-py2.7.egg/pkg_resources.py", line 1688, in find_on_path
    return ()
  File "/usr/lib/python2.7/site-packages/setuptools-0.6c11-py2.7.egg/pkg_resources.py", line 1835, in _normalize_cached
    File "/usr/lib/python2.7/site-packages/setuptools-0.6c11-py2.7.egg/pkg_resources.py", line 1829, in normalize_path
    register_namespace_handler(object.null_ns_handler)
  File "/usr/lib64/python2.7/posixpath.py", line 368, in realpath
    return abspath(path)
  File "/usr/lib64/python2.7/posixpath.py", line 356, in abspath
    cwd = os.getcwd()
OSError: [Errno 2] No such file or directory
```

### Cause Analysis

After the service configuration is modified and the service is restarted, some directory structures of the service, such as the etc directory, are deleted and recreated. If the directory is etc or its subdirectory before the service is restarted, some system variables or parameters cannot be found when impala-shell is

executed in the original directory because the directory is recreated after the service is restarted. As a result, `impala-shell` fails to be connected.

## Procedure

Switch to any existing directory and reconnect to `impala-shell`.

## 15.2 Failed to Create a Kudu Table

### Issue

An error occurs when a user creates a Kudu table.

### Symptom

A cluster is created. When a table is created, the following error message is displayed:

```
[Cloudera]ImpalaJDBCdriver ERROR processing query/statement. Error Code: 0, SQL state:  
TStatus(statusCode:ERROR_STATUS, sqlState:HY000, errorMessage:AnalysisException: Table property  
'kudu.master_addresses' is required when the impalad startup flag -kudu_master_hosts is not used."
```

### Cause Analysis

`kudu.master_addresses` is not specified in the Impala SQL statement.

## Procedure

Specify `kudu.master_addresses` when creating a Kudu table. For details, see [Using Impala to Operate Kudu](#).

## 15.3 Installing Python2 on the Impala Client

### Issue

Error information similar to the following is displayed when a user runs the Impala client.

```
[root@node-master1avIy ~]# impala-shell -i 192.168.128.49:21000  
File "/opt/client/Impala/impala/shell/impala_shell.py", line 1675  
except Exception, e:  
    ^  
SyntaxError: invalid syntax  
[root@node-master1avIy ~]#
```

### Cause Analysis

The latest MRS cluster uses EulerOS 2.9 or later, which provides only Python 3. However, the Impala client is implemented based on Python 2 and is incompatible with some syntax of Python 3. As a result, an error occurs when the Impala client is running. You can manually install Python 2 to solve this problem.

## Procedure

- Step 1** Log in to the Impala node as user **root** and run the following command to check its Python version:

```
python --version
```

```
[root@node-master2Jg0Y ~]# python --version  
Python 3.7.4
```

- Step 2** Run the **yum install make** command to check whether yum is available.

- If the following error is reported, the yum configuration is incorrect. Go to [Step 3](#).

```
[root@node-master2Jg0Y ~]# yum install make  
Error: There are no enabled repositories in "/etc/yum.repos.d", "/etc/yum/repos.d", "/etc/distro.repos.d".
```

- If no error is reported, go to [Step 4](#).

- Step 3** Run the **cat /etc/yum.repos.d/EulerOS-base.repo** command to check whether the yum source matches the system version. If they do not match, modify them.

Before modification

```
[root@node-master1avIy ~]# cat /etc/yum.repos.d/EulerOS-base.repo  
[base]  
name=EulerOS-2.0SP2 base  
baseurl=http://mirrors.my[REDACTED].com/euler/ict/site-euleros/euleros/repo/yum/2.2/os/x86_64/  
enabled=1  
gpgcheck=1  
gpgkey=http://mirrors.my[REDACTED].com/euler/ict/site-euleros/euleros/repo/yum/2.2/os/RPM-GPG-KEY-EulerOS  
[root@node-master1avIy ~]# uname -a  
Linux node-master1avIy.mrs-mq7v.com 4.18.0-147.5.1.6.h541.eulerosv2r9.x86_64 #1 SMP Wed Aug 4 02:30:13 UTC  
x86_64 GNU/Linux
```

After modification

```
[base]  
name=EulerOS-2.0SP9 base  
baseurl=http://mirrors.my[REDACTED].com/euler/ict/site-euleros/euleros/repo/yum/2.9/os/x86_64/  
enabled=1  
gpgcheck=1  
gpgkey=http://mirrors.my[REDACTED].com/euler/ict/site-euleros/euleros/repo/yum/2.9/os/RPM-GPG-KEY-EulerOS
```

- Step 4** Run the following command to check for the software whose name starts with **python2** in the yum source:

```
yum list python2*
```

```
[root@node-master2Jg0Y ~]# yum list python2*  
Last metadata expiration check: 0:02:36 ago on Thu 16 Dec 2021 10:05:52 AM CST.  
Available Packages  
python2.x86_64 2.7.16-16.eulerosv2r9  
python2-debug.x86_64 2.7.16-16.eulerosv2r9  
python2-devel.x86_64 2.7.16-16.eulerosv2r9  
python2-help.noarch 2.7.16-16.eulerosv2r9  
python2-pip.noarch 18.0-13.h2.eulerosv2r9  
python2-setuptools.noarch 40.4.3-4.h1.eulerosv2r9  
python2-tkinter.x86_64 2.7.16-16.eulerosv2r9  
python2-tools.x86_64 2.7.16-16.eulerosv2r9
```

- Step 5** Run the following command to install Python 2:

```
yum install python2
```

```
[root@node-master2Jg0Y ~]# yum install python2
Last metadata expiration check: 0:00:48 ago on Thu 16 Dec 2021 10:05:52 AM CST.
Error:
Problem: problem with installed package python3-unversioned-command-3.7.4-7.h29.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h29.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h11.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h13.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h15.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h18.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h33.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- package python3-unversioned-command-3.7.4-7.h38.eulerosv2r9.x86_64 conflicts with python2 provided by python2-2.7.16-16.eulerosv2r9.x86_64
- conflicting requests
(tr try to add '--allowrasing' to command line to replace conflicting packages or '--skip-broken' to skip uninstalleable packages or '--nobest' to use not only best candidate packages)
```

Python 3 has been installed in the current system. If you directly install Python 2, a conflict message is displayed.

You can select **--allowrasing** or **--skip-broken** for the installation. For example:

**yum install python2 --skip-broken**

```
[root@node-master2Jg0Y ~]# yum install python2 --skip-broken
Last metadata expiration check: 0:34:08 ago on Thu 16 Dec 2021 10:05:52 AM CST.
Dependencies resolved.
=====
Package                        Architecture      Version           Repository        Size
=====
Installing:
python2                        x86_64            2.7.16-16.eulerosv2r9    base              6.4 M
Installing dependencies:
libXft                          x86_64            2.3.2-13.eulerosv2r9    base              41 k
=====
```

After the installation is complete, the Python version is automatically changed to python2, as shown in the following figure.

```
Installed:
libXft-2.3.2-13.eulerosv2r9.x86_64          libXrender-0.9.10-10.eulerosv2r9.x86_64
python2-2.7.16-16.eulerosv2r9.x86_64      python2-debug-2.7.16-16.eulerosv2r9.x86_64
python2-devel-2.7.16-16.eulerosv2r9.x86_64 python2-help-2.7.16-16.eulerosv2r9.noarch
python2-setuptools-40.4.3-4.h1.eulerosv2r9.noarch python2-tkinter-2.7.16-16.eulerosv2r9.x86_64
python2-tools-2.7.16-16.eulerosv2r9.x86_64 python3-rpm-generators-9-1.eulerosv2r9.noarch
tk-1:8.6.8-5.eulerosv2r9.x86_64

Complete!
[root@node-master2Jg0Y ~]# python --version
Python 2.7.16
```

If Python 2 is installed successfully but the displayed Python version is incorrect, run the following command to create the **/usr/bin/python** soft link for **/usr/bin/python2**:

```
ln -s /usr/bin/python2 /usr/bin/python
```

**Step 6** Verify that the Impala client is available.

```
[root@node-master1avIy ~]# impala-shell -i 192.168.128.49:21000
Starting Impala Shell without Kerberos authentication
Opened TCP connection to 192.168.128.49:21000
Connected to 192.168.128.49:21000
Server version: impalad version 3.4.0-RELEASE RELEASE (build eebadd34c1563cbf5825a4e4d361e7b3601f9827)
*****
Welcome to the Impala shell.
(Impala Shell v3.4.0-RELEASE (eebadd3) built on Thu Nov  4 11:29:54 CST 2021)

After running a query, type SUMMARY to see a summary of where time was spent.
*****
[192.168.128.49:21000] default> show databases;
Query: show databases
+-----+-----+
| name          | comment                                     |
+-----+-----+
| _impala_builtins | System database for Impala builtin functions |
| default        | Default Hive database                       |
+-----+-----+
Fetched 2 row(s) in 0.16s
[192.168.128.49:21000] default>
```

----End



# 16 Using Kafka

---

## 16.1 An Error Is Reported When the Kafka Client Is Run to Obtain Topics

### Symptom

The following error is reported when the Kafka client is run to obtain topics:

```
ERROR org.apache.kafka.common.errors.InvalidReplicationFactorException: Replication factor: 2 larger than available brokers: 0.
```

### Possible Cause

The variable for obtaining the ZooKeeper address is incorrect due to special characters.

### Procedure

- Step 1** Log in to any Master node.
  - Step 2** Run the `cat Client installation directory/Kafka/kafka/config/server.properties | grep '^zookeeper.connect ='` command to check the variable of the ZooKeeper address.
  - Step 3** Run Kafka again to obtain the topic. Do not add any character to the variables obtained in [Step 2](#).
- End

## 16.2 Using Python3.x to Connect to Kafka in a Security Cluster

### Issue

The user does not know how to connect to a Kafka cluster with Kerberos authentication enabled in the Python3.x environment.

## Symptom

The user needs an operation guide to connect to the Kafka cluster with Kerberos authentication enabled in the Python3.x environment.

## Procedure

- Step 1** Log in to the master node and run the following command to configure the Huawei Cloud EulerOS image source:

```
wget http://mirrors.myhuaweicloud.com/repo/mirrors_source.sh && sh mirrors_source.sh
```

- Step 2** Run the following commands to compile Python3.x:

```
yum groupinstall "Development tools" -y
yum -y install zlib zlib-devel
yum -y install bzip2 bzip2-devel
yum -y install ncurses ncurses-devel
yum -y install readline readline-devel
yum -y install openssl openssl-devel
yum -y install openssl-static
yum -y install xz lzma xz-devel
yum -y install sqlite sqlite-devel
yum -y install gdbm gdbm-devel
yum -y install tk tk-devel
yum -y install libffi libffi-devel
```

- Step 3** After the compilation is successful, run the following commands to download and decompress the .tgz package of Python3.x:

```
wget https://www.python.org/ftp/python/3.6.7/Python-3.6.7.tgz
tar -zxvf Python-3.6.7.tgz
cd Python-3.6.7
```

### NOTE

You can also download the .tgz package of Python3.x from the Python official website. Python-3.6.X is recommended. In version 3.7, the **take** function of RDD cannot be used.

- Step 4** Run the following commands to configure Python3.x, and compile and install it in the `/opt/Bigdata/python3` directory:

```
./configure --prefix=/opt/Bigdata/python3 --enable-shared CFLAGS=-fPIC
make && make install
```

 NOTE

The installation directory can be customized.

**Step 5** Run the following commands to configure Python3.x variables:

```
echo "/opt/Bigdata/python3/lib" >> /etc/ld.so.conf
```

```
ldconfig
```

```
ln -s /opt/Bigdata/python3/bin/python3 /usr/bin/python3
```

```
ln -s /opt/Bigdata/python3/bin/pip3 /usr/bin/pip3
```

 NOTE

The variable directory must be the same as the installation directory specified in [Step 4](#).

**Step 6** After the configuration is successful, run the following commands to install Kafka in the Python3.x environment:

```
cp /usr/include/gssapi/* /home/omm/kerberos/include/gssapi/
```

```
pip3 install kafka-python
```

```
pip3 install gssapi
```

**Step 7** After the installation is successful, run the following command to configure environment variables:

```
source Client installation directory/bigdata_env
```

**Step 8** Run the following command to authenticate the current user:

```
kinit Kafka user
```

 NOTE

The Kafka user is the one who logs in to Manager. This user must have the permissions of the Kafka user group.

**Step 9** Run the Python3.x script.

Sample script:

```
producer:
from kafka import KafkaProducer
producer = KafkaProducer(bootstrap_servers=["broker_ip:21007"],
security_protocol="SASL_PLAINTEXT",
saslm_mechanism="GSSAPI",
saslm_kerberos_service_name="kafka",
saslm_kerberos_domain_name="hadoop.hadoop.com")
for _ in range(100):
response = producer.send("test-topic", b"testmessage")
result = response.get(timeout=50)
print(result)

consumer:
from kafka import KafkaConsumer
consumer = KafkaConsumer("test-topic",
bootstrap_servers=["broker_ip:21007"],
group_id="test-group",
enable_auto_commit="true",
security_protocol="SASL_PLAINTEXT",
saslm_mechanism="GSSAPI",
```

```
sasl_kerberos_service_name="kafka",  
sasl_kerberos_domain_name="hadoop.hadoop.com")  
for message in consumer:  
    print(message)
```

----End

## 16.3 Flume Normally Connects to Kafka but Fails to Send Messages

### Symptom

An MRS cluster is installed, and ZooKeeper, Flume, and Kafka are installed in the cluster.

Flume fails to send data to Kafka.

### Possible Causes

1. The Kafka service is abnormal.
2. The IP address for Flume to connect to Kafka is incorrect.
3. The size of the message sent from Flume to Kafka exceeds the upper limit.

### Cause Analysis

The possible reasons why Flume fails to send data to Kafka may be related to Flume or Kafka.

1. Check the Kafka service status and monitoring metrics on Manager.
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.

2. Check the Flume log. The log contains **MessageSizeTooLargeException** information, as shown in the following:

```
2016-02-26 14:55:19,126 | WARN | [SinkRunner-PollingRunner-DefaultSinkProcessor] | Produce  
request with correlation id 349829 failed due to [LOG,7]:  
kafka.common.MessageSizeTooLargeException | kafka.utils.Logging$class.warn(Logging.scala:83)
```

The exception shows that the size of data written to Kafka by Flume exceeds the maximum message size specified by Kafka.

3. Check the maximum message size specified by Kafka on Manager.
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka > Configurations**.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka > Configurations**.

On the page that is displayed, set **Type** to **All**. All Kafka configurations are displayed. Enter **message.max.bytes** in the **Search** text box to search.

In MRS, the maximum size of a message that can be received by the Kafka server is 977 KB (1000012 bytes) by default.

## Solution

After confirmation with the customer, data sent by Flume contains messages over 1 MB. Adjust parameters on Kafka to enable the messages to be written to Kafka.

- Step 1** Set **message.max.bytes** to a value that is larger than the current maximum size of the message to be written so that Kafka can receive all messages.
- Step 2** Set **replica.fetch.max.bytes** to a value that is equal to or larger than the value of **message.max.bytes** so that replicas of partitions on different Brokers can be synchronized to all messages.
- MRS Manager: Log in to MRS Manager and choose **Services > Kafka > Configurations**.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka > Configurations**.

On the page that is displayed, set **Type** to **All**. All Kafka configurations are displayed. Enter **replica.fetch.max.bytes** in the **Search** text box to search.

- Step 3** Click **Save** and restart the Kafka service to make Kafka configurations take effect.
- Step 4** Set **fetch.message.max.bytes** to a value that is equal to or larger than the value of **message.max.bytes** for Consumer service applications to ensure that Consumers can consume all messages.

----End

## 16.4 Producer Fails to Send Data and Error Message "NullPointerException" Is Displayed

### Symptom

An MRS cluster has ZooKeeper and Kafka installed.

When the Producer sends data to Kafka, the client throws "NullPointerException".

### Possible Causes

1. The Kafka service is abnormal.
2. The **jaas** and **keytab** files of the Producer on the client are incorrect.

### Cause Analysis

The possible causes may be related to Producer or Kafka.

1. Check the Kafka service status and monitoring metrics on Manager.
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.

2. Check the Producer client log. The log contains "NullPointerException", as shown in [Figure 16-1](#).

**Figure 16-1** Producer client log

```
[2016-12-06 02:04:05,906]-[schedule-C50D0717-4643-4D4E-9D6E-B940E4BD7159]-[kafka-producer-network-thread |
SZX1000161910-10.21.219.222-bigdata-producer-5]-[1005]-[org.apache.kafka.clients.producer.internals.Sender.run
thread:
java.lang.NullPointerException
    at org.apache.kafka.common.network.Selector.pollSelectionKeys(Selector.java:302)
    at org.apache.kafka.common.network.Selector.poll(Selector.java:283)
    at org.apache.kafka.clients.NetworkClient.poll(NetworkClient.java:260)
    at org.apache.kafka.clients.producer.internals.Sender.run(Sender.java:229)
    at org.apache.kafka.clients.producer.internals.Sender.run(Sender.java:134)
    at java.lang.Thread.run(Thread.java:745)
[2016-12-06 02:04:05,921]-[schedule-C50D0717-4643-4D4E-9D6E-B940E4BD7159]-[kafka-producer-network-thread |
SZX1000161910-10.21.219.222-bigdata-producer-3]-[1005]-[org.apache.kafka.clients.producer.internals.Sender.run
thread:
java.lang.NullPointerException
    at org.apache.kafka.common.network.Selector.pollSelectionKeys(Selector.java:302)
    at org.apache.kafka.common.network.Selector.poll(Selector.java:283)
    at org.apache.kafka.clients.NetworkClient.poll(NetworkClient.java:260)
    at org.apache.kafka.clients.producer.internals.Sender.run(Sender.java:229)
    at org.apache.kafka.clients.producer.internals.Sender.run(Sender.java:134)
    at java.lang.Thread.run(Thread.java:745)
```

Alternatively, the log contains only "NullPointerException" but no stack information. The problem is caused by JDK self-protection. If much information is printed for the same stack, the JDK self-protection is triggered and stack information is no longer printed, as shown in [Figure 16-2](#).

**Figure 16-2** Error information

```
[2016-11-23 04:06:53,973] [kafka-producer-network-thread | producer-1] [ERROR] [org.apache.kafka.clients.producer.internals.Sender] (run:130)- Uncaught error in kafka producer I/O thread:
java.lang.NullPointerException
[2016-11-23 04:06:53,973] [kafka-producer-network-thread | producer-1] [ERROR] [org.apache.kafka.clients.producer.internals.Sender] (run:130)- Uncaught error in kafka producer I/O thread:
java.lang.NullPointerException
[2016-11-23 04:06:53,973] [kafka-producer-network-thread | producer-1] [ERROR] [org.apache.kafka.clients.producer.internals.Sender] (run:130)- Uncaught error in kafka producer I/O thread:
java.lang.NullPointerException
[2016-11-23 04:06:53,973] [kafka-producer-network-thread | producer-1] [ERROR] [org.apache.kafka.clients.producer.internals.Sender] (run:130)- Uncaught error in kafka producer I/O thread:
java.lang.NullPointerException
```

3. Check the Producer client log. Error information "Failed to configure SaslClientAuthenticator" is displayed, as shown in [Figure 16-3](#).

**Figure 16-3** Error log

```
Caused by: org.apache.kafka.common.KafkaException: Failed to configure SaslClientAuthenticator
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator.configure(SaslClientAuthenticator.java:96)
at org.apache.kafka.common.network.SaslChannelBuilder.buildChannel(SaslChannelBuilder.java:89)
... 9 more
Caused by: org.apache.kafka.common.KafkaException: Failed to create SaslClient
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator.createSaslClient(SaslClientAuthenticator.java:112)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator.configure(SaslClientAuthenticator.java:94)
... 10 more
Caused by: javax.security.sasl.SaslException: PLAIN: authorization ID and password must be specified
at com.sun.security.sasl.PlainClient.<init>(PlainClient.java:58)
at com.sun.security.sasl.ClientFactoryImpl.createSaslClient(ClientFactoryImpl.java:97)
at javax.security.sasl.Sasl.createSaslClient(Sasl.java:384)
at com.ibm.messagehub.login.MessageHubSaslClientFactory.createSaslClient(MessageHubSaslClientFactory.java:77)
at javax.security.sasl.Sasl.createSaslClient(Sasl.java:384)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator$1.run(SaslClientAuthenticator.java:107)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator$1.run(SaslClientAuthenticator.java:102)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator.createSaslClient(SaslClientAuthenticator.java:102)
... 11 more
```

4. The authentication failure causes the failure to create the KafkaChannel. The KafkaChannel obtained through the **channel(key)** method is empty and "NullPointerException" is excessively printed. According to the preceding log, the authentication fails due to an incorrect password which does not match the username.
5. Check the **jaas** and **keytab** files. The **principal** is set to **stream** in the **jaas** file.

**Figure 16-4** Checking the jaas file

```
kafkaClient {  
com.sun.security.auth.module.Krb5LoginModule required  
debug=false  
keyTab="/opt/client/user.keytab"  
useTicketCache=false  
storeKey=true  
principal="stream@HADOOP.COM"  
useKeyTab=true;  
};
```

The **principal** is set to **zmk\_kafka** in the **user.keytab** file.

**Figure 16-5** Checking the user.keytab file

```
[root@8-5-148-6 client]# klist -kt user.keytab  
Keytab name: FILE:user.keytab  
KVNO Timestamp Principal  
-----  
1 12/19/16 16:28:17 zmk_kafka@HADOOP.COM  
1 12/19/16 16:28:17 zmk_kafka@HADOOP.COM
```

The **principal** in the **jaas** file is inconsistent with that in the **user.keytab** file.

The application automatically and periodically updates the **jaas** file. However, when two processes of the application update the **jaas** file, one process writes a correct **principal** whereas the other process writes an incorrect one. As a result, the application is abnormal sometimes.

## Procedure

**Step 1** Modify the **jaas** file to ensure that its **principal** exists in the **keytab** file.

----End

## 16.5 Producer Fails to Send Data and Error Message "TOPIC\_AUTHORIZATION\_FAILED" Is Displayed

### Symptom

An MRS cluster is installed, and ZooKeeper and Kafka are installed in the cluster.

When Producer sends data to Kafka, the client reports the error "TOPIC\_AUTHORIZATION\_FAILED."

### Possible Causes

1. The Kafka service is abnormal.
2. The Producer client adopts non-security access and access is disabled on the server.
3. The Producer client adopts non-security access and ACL is set for Kafka topics.

## Cause Analysis

The possible reasons why Producer fails to send data to Kafka may be related to Producer or Kafka.

1. Check the Kafka service status:
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.

2. Check the Producer client logs. The logs contain the error information "TOPIC\_AUTHORIZATION\_FAILED."

```
[root@10-10-144-2 client]# kafka-console-producer.sh --broker-list 10.5.144.2:9092 --topic test
1
[2017-01-24 16:58:36,671] WARN Error while fetching metadata with correlation id 0 :
{test=TOPIC_AUTHORIZATION_FAILED} (org.apache.kafka.clients.NetworkClient)
[2017-01-24 16:58:36,672] ERROR Error when sending message to topic test with key: null, value: 1
bytes with error: Not authorized to access topics: [test]
(org.apache.kafka.clients.producer.internals.ErrorLoggingCallback)
```

Producer accesses Kafka using port 9092, which is a non-security port.

3. On Manager, check the current Kafka cluster configuration. It is found that the customized configuration **allow.everyone.if.no.acl.found=false** is not configured.

- MRS Manager: Log in to MRS Manager and choose **Services > Kafka > Configurations**.
- FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka > Configurations**.

4. If ACL is set to **false**, port 9092 cannot be used for access.

5. Check the Producer client logs. The logs contain the error information "TOPIC\_AUTHORIZATION\_FAILED."

```
[root@10-10-144-2 client]# kafka-console-producer.sh --broker-list 10.5.144.2:21005 --topic test_acl
1
[2017-01-25 11:09:40,012] WARN Error while fetching metadata with correlation id 0 :
{test_acl=TOPIC_AUTHORIZATION_FAILED} (org.apache.kafka.clients.NetworkClient)
[2017-01-25 11:09:40,013] ERROR Error when sending message to topic test_acl with key: null, value:
1 bytes with error: Not authorized to access topics: [test_acl]
(org.apache.kafka.clients.producer.internals.ErrorLoggingCallback)
[2017-01-25 11:14:40,010] WARN Error while fetching metadata with correlation id 1 :
{test_acl=TOPIC_AUTHORIZATION_FAILED} (org.apache.kafka.clients.NetworkClient)
```

Producer accesses Kafka using port 21005, which is a non-security port.

6. Run the client command to check the ACL permission of the topic.

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:24002/
kafka --list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
User:test_user has Allow permission for operations: Write from hosts: *
```

If ACL is set for the topic, port 9092 cannot be used for access.

7. Check the Producer client logs. The logs contain the error information "TOPIC\_AUTHORIZATION\_FAILED."

```
[root@10-10-144-2 client]# kafka-console-producer.sh --broker-list 10.5.144.2:21007 --topic topic_acl
--producer.config /opt/client/Kafka/kafka/config/producer.properties
1
[2017-01-25 12:43:58,506] WARN Error while fetching metadata with correlation id 0 :
{topic_acl=TOPIC_AUTHORIZATION_FAILED} (org.apache.kafka.clients.NetworkClient)
```



```
[2017-01-25 12:43:58,507] ERROR Error when sending message to topic topic_acl with key: null,
value: 1 bytes with error: Not authorized to access topics: [topic_acl]
(org.apache.kafka.clients.producer.internals.ErrorLoggingCallback)
```

Producer uses port 21007 to access Kafka.

8. Run the client command **klist** to query the current authenticated user.

```
[root@10-10-144-2 client]# klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: test@HADOOP.COM

Valid starting Expires Service principal
01/25/17 11:06:48 01/26/17 11:06:45 krbtgt/HADOOP.COM@HADOOP.COM
```

The **test** user is used in this example.

9. Run the client command to check the ACL permission of the topic.

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/
kafka --list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
User:test_user has Allow permission for operations: Write from hosts: *
```

After ACL is set for the topic, user **test\_user** has Producer permission. User **test** has no permission to perform Producer operations.

For details about the solution, see [Step 2](#).

10. Log in to Kafka Broker using SSH.

Run the **cd /var/log/Bigdata/kafka/broker** command to go to the log directory.

Check the **kafka-authorizer.log** file. It shows that the user does not belong to the **kafka** or **kafkaadmin** group.

```
2017-01-25 13:26:33,648 | INFO | [kafka-request-handler-0] | The principal is test, belongs to Group :
[hadoop, ficommon] | kafka.authorizer.logger (SimpleAclAuthorizer.scala:169)
2017-01-25 13:26:33,648 | WARN | [kafka-request-handler-0] | The user is not belongs to kafka or
kafkaadmin group, authorize failed! | kafka.authorizer.logger (SimpleAclAuthorizer.scala:170)
```

For details about the solution, see [Step 3](#).

## Solution

**Step 1** Set **allow.everyone.if.no.acl.found** to **true** and restart the Kafka service.

**Step 2** Use the account with permission for login.

Example:

```
kinit test_user
```

Alternatively, grant the user with related permission.

### NOTE

This operation must be performed by the Kafka administrator (belonging to the **kafkaadmin** group).

Example:

```
kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/kafka
--topic topic_acl --producer --add --allow-principal User:test
```

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=8.5.144.2:2181/kafka --
list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
```

```
User:test_user has Allow permission for operations: Write from hosts: *
User:test has Allow permission for operations: Describe from hosts: *
User:test has Allow permission for operations: Write from hosts: *
```

**Step 3** Add the user to the **kafka** or **kafkaadmin** group.

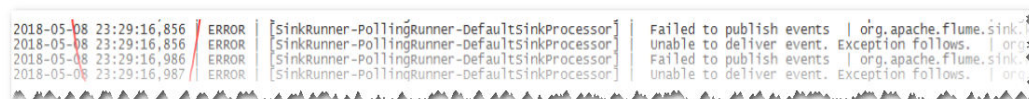
----End

## 16.6 Producer Occasionally Fails to Send Data and the Log Displays "Too many open files in system"

### Symptom

When Producer sends data to Kafka, it is found that the client fails to send data.

**Figure 16-6** Producer fails to send data.



```
2018-05-08 23:29:16,856 ERROR [SinkRunner-PollingRunner-DefaultSinkProcessor] Failed to publish events org.apache.flume.sink...
2018-05-08 23:29:16,856 ERROR [SinkRunner-PollingRunner-DefaultSinkProcessor] Unable to deliver event. Exception follows. org...
2018-05-08 23:29:16,986 ERROR [SinkRunner-PollingRunner-DefaultSinkProcessor] Failed to publish events org.apache.flume.sink...
2018-05-08 23:29:16,987 ERROR [SinkRunner-PollingRunner-DefaultSinkProcessor] Unable to deliver event. Exception follows. org...
```

### Possible Causes

1. The Kafka service is abnormal.
2. The network is abnormal.
3. The Kafka topic is abnormal.

### Cause Analysis

1. Check the Kafka service status:
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.

2. View the error topic information in the SparkStreaming log.

Run the Kafka commands to obtain the topic assignment information and copy synchronization information, and check the return result.

**kafka-topics.sh --describe --zookeeper <zk\_host:port/chroot>**

As shown in [Figure 16-7](#), the topic status is normal. All partitions have normal leader information.

Figure 16-7 Topic status

Topic: STK6	Partition: 36	Leader: 3	Replicas: 3,5	Isr: 3,5
Topic: STK6	Partition: 37	Leader: 4	Replicas: 4,6	Isr: 4,6
Topic: STK6	Partition: 38	Leader: 5	Replicas: 5,7	Isr: 5,7
Topic: STK6	Partition: 39	Leader: 6	Replicas: 6,8	Isr: 6,8
Topic: STK6	Partition: 40	Leader: 7	Replicas: 7,9	Isr: 7,9
Topic: STK6	Partition: 41	Leader: 8	Replicas: 8,1	Isr: 8,1
Topic: STK6	Partition: 42	Leader: 9	Replicas: 9,2	Isr: 9,2
Topic: STK6	Partition: 43	Leader: 1	Replicas: 1,3	Isr: 3,1
Topic: STK6	Partition: 44	Leader: 2	Replicas: 2,4	Isr: 2,4
Topic: STK6	Partition: 45	Leader: 3	Replicas: 3,6	Isr: 3,6
Topic: STK6	Partition: 46	Leader: 4	Replicas: 4,7	Isr: 4,7
Topic: STK6	Partition: 47	Leader: 5	Replicas: 5,8	Isr: 5
Topic: STK6	Partition: 48	Leader: 6	Replicas: 6,9	Isr: 6,9
Topic: STK6	Partition: 49	Leader: 7	Replicas: 7,1	Isr: 7,1
Topic: STK6	Partition: 50	Leader: 8	Replicas: 8,2	Isr: 2,8
Topic: STK6	Partition: 51	Leader: 9	Replicas: 9,3	Isr: 9,3
Topic: STK6	Partition: 52	Leader: 1	Replicas: 1,4	Isr: 4,1
Topic: STK6	Partition: 53	Leader: 2	Replicas: 2,5	Isr: 5,2
Topic: STK6	Partition: 54	Leader: 3	Replicas: 3,7	Isr: 3,7
Topic: STK6	Partition: 55	Leader: 4	Replicas: 4,8	Isr: 4,8
Topic: STK6	Partition: 56	Leader: 5	Replicas: 5,9	Isr: 5,9
Topic: STK6	Partition: 57	Leader: 6	Replicas: 6,1	Isr: 6,1
Topic: STK6	Partition: 58	Leader: 7	Replicas: 7,2	Isr: 2,7
Topic: STK6	Partition: 59	Leader: 8	Replicas: 8,3	Isr: 3,8

- Run the **telnet** command to check whether the Kafka can be connected.

**telnet** *Kafka service IP address* *Kafka service port*

If Telnet fails, check the network security group and ACL.

- Log in to Kafka Broker using SSH.

Run the **cd /var/log/Bigdata/kafka/broker** command to go to the log directory.

Check on **server.log** indicates that the error message is displayed in the log shown in the following figure.

Figure 16-8 Log exception

```
2018-05-08 23:05:00,061 | ERROR | [kafka-socket-acceptor-PLAINTEXT-21005] | Error while accepting connection | kafka.network.Acceptor.accept(SocketServer.scala:336)
java.io.IOException: Too many open files in system
    at sun.nio.ch.ServerSocketChannelImpl.accept0(Native Method)
    at sun.nio.ch.ServerSocketChannelImpl.accept(ServerSocketChannelImpl.java:422)
    at sun.nio.ch.ServerSocketChannelImpl.accept(ServerSocketChannelImpl.java:250)
    at kafka.network.Acceptor.accept(SocketServer.scala:336)
```

- Output of the **lsdf** command used to check the handle usage of the Kafka process on the current node shows that the number of handles used by the Kafka process reaches 470,000.

Figure 16-9 Handles

```
omm@lf2-bi-sparkstream-71-24-8:/var/log/Bigdata/kafka/broker> jps
24338 Kafka
14630 MetricAgentMain
30713 NodeAgent
46973 Jps
omm@lf2-bi-sparkstream-71-24-8:/var/log/Bigdata/kafka/broker> lsdf -p 24383|wc
0
omm@lf2-bi-sparkstream-71-24-8:/var/log/Bigdata/kafka/broker> lsdf -p 24338|wc
473282
```

6. Check the service code. It is found that new Producer objects are frequently created and are not closed normally.

## Solution

**Step 1** Stop the current application to ensure that the number of handles on the server does not increase sharply, which affects the normal running of services.

**Step 2** Optimize the application code to resolve the handle leakage problem.

Suggestion: Use one Producer object globally. After the use is complete, call the Close interface to close the handle.

----End

## 16.7 Consumer Is Initialized Successfully, but the Specified Topic Message Cannot Be Obtained from Kafka

### Symptom

An MRS cluster is installed, and ZooKeeper, Flume, Kafka, Storm, and Spark are installed in the cluster.

The customer cannot consume any data using Storm, Spark, Flume or self-programmed Consumer code to consume messages of the specified Kafka topic.

### Possible Causes

1. The Kafka service is abnormal.
2. The IP address for ZooKeeper connection is incorrectly set.
3. "ConsumerRebalanceFailedException" is thrown.
4. "ClosedChannelException" caused by network problems is thrown.

### Cause Analysis

Storm, Spark, Flume or user-defined Consumer code can be called Consumer.

1. Check the Kafka service status:
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.

2. Check whether data can be normally consumed through the Kafka client.

Suppose the client has been installed in the `/opt/client` directory, `test` is the topic name to be consumed, and the IP address of ZooKeeper is 192.168.234.231.

```
cd /opt/client
source bigdata_env
```

**kinit admin**

```
kafka-topics.sh --zookeeper 192.168.234.231:2181/kafka --describe --topic testkafka-console-consumer.sh --topic test --zookeeper 192.168.234.231:2181/kafka --from-beginning
```

If data can be consumed, the cluster service is running properly.

3. Check Consumer configurations. The IP address for connecting to ZooKeeper is incorrect.

- **Flume**

```
server.sources.Source02.type=org.apache.flume.source.kafka.KafkaSource
```

```
server.sources.Source02.zookeeperConnect=192.168.234.231:2181
server.sources.Source02.topic = test
server.sources.Source02.groupId = test_01
```

- **Spark**

```
val zkQuorum = "192.168.234.231:2181"
```

- **Storm**

```
BrokerHosts brokerHosts = new ZKHosts("192.168.234.231:2181");
```

- **Consumer API**

```
zookeeper.connect="192.168.234.231:2181"
```

On MRS Manager, the root path of ZNode where Kafka is stored on ZooKeeper is **/kafka**, which is differentiated from the open source. The address for Kafka to connect to ZooKeeper is **192.168.234.231:2181/kafka**.

However, the address for Consumer to connect to ZooKeeper is **192.168.234.231:2181**. Therefore, topic information about Kafka cannot be correctly obtained.

For details about the solution, see [Step 1](#).

4. Check Consumer logs. The logs contain "ConsumerRebalanceFailedException".

```
2016-02-03 15:55:32,557 | ERROR | [ZkClient-EventThread-75- 192.168.234.231:2181/kafka] | Error handling event ZkEvent[New session event sent to kafka.consumer.ZookeeperConsumerConnector $ZKSessionExpireListener@34b41dfe] | org.I0ltec.zkclient.ZkEventThread.run(ZkEventThread.java:77)
kafka.common.ConsumerRebalanceFailedException: pc-zjqbetl86-1454482884879-2ec95ed3 can't rebalance after 4 retries
at kafka.consumer.ZookeeperConsumerConnector
$ZKRebalancerListener.syncedRebalance(ZookeeperConsumerConnector.scala:633)
at kafka.consumer.ZookeeperConsumerConnector
$ZKSessionExpireListener.handleNewSession(ZookeeperConsumerConnector.scala:487)
at org.I0ltec.zkclient.ZkClient$4.run(ZkClient.java:472)
at org.I0ltec.zkclient.ZkEventThread.run(ZkEventThread.java:71)
```

The exception shows that the current Consumer does not complete rebalance within the specified retry times. As a result, Kafka Topic-Partition is not allocated to Consumer and Consumer cannot consume messages.

For details about the solution, see [Step 3](#).

5. Check Consumer logs. The logs contain "Fetching topic metadata with correlation id 0 for topics [Set(test)] from broker [id:26,host:192-168-234-231,port:9092] failed" and "ClosedChannelException".

```
[2016-03-04 03:33:53,047] INFO Fetching metadata from broker id:26,host:
192-168-234-231,port:9092 with correlation id 0 for 1 topic(s) Set(test) (kafka.client.ClientUtils$)
[2016-03-04 03:33:55,614] INFO Connected to 192-168-234-231:21005 for producing
(kafka.producer.SyncProducer)
[2016-03-04 03:33:55,614] INFO Disconnecting from 192-168-234-231:21005
(kafka.producer.SyncProducer)
[2016-03-04 03:33:55,615] WARN Fetching topic metadata with correlation id 0 for topics [Set(test)]
from broker [id:26,host: 192-168-234-231,port:21005] failed (kafka.client.ClientUtils$)
java.nio.channels.ClosedChannelException
at kafka.network.BlockingChannel.send(BlockingChannel.scala:100)
at kafka.producer.SyncProducer.liftedTree1$1(SyncProducer.scala:73)
at kafka.producer.SyncProducer.kafka$producer$SyncProducer$$doSend(SyncProducer.scala:72)
at kafka.producer.SyncProducer.send(SyncProducer.scala:113)
at kafka.client.ClientUtils$.fetchTopicMetadata(ClientUtils.scala:58)
```

```
at kafka.client.ClientUtils$.fetchTopicMetadata(ClientUtils.scala:93)
at kafka.consumer.ConsumerFetcherManager
$LeaderFinderThread.doWork(ConsumerFetcherManager.scala:66)
at kafka.utils.ShutdownableThread.run(ShutdownableThread.scala:60)
[2016-03-04 03:33:55,615] INFO Disconnecting from 192-168-234-231:21005
(kafka.producer.SyncProducer)
```

The exception shows that the current Consumer cannot obtain metadata from the Kafka Broker 192-168-234-231 node and cannot connect to the correct Broker for obtaining messages.

6. Check the network conditions. If the network is normal, check whether mapping between the host and the IP address is configured.

- Linux

Run the `cat /etc/hosts` command.

```
127.0.0.1    localhost
            |
192.168.0.131 192-168-0-131
192.168.0.51  192-168-0-51
192.168.0.122 192-168-0-122
```

- Windows

Open `C:\Windows\System32\drivers\etc\hosts`.

```
# For example:
#
#       192.168.94.97    rhino.acme.com       # source server
#       192.168.63.10   x.acme.com           # x client host

# localhost name resolution is handled within DNS itself.
#       127.0.0.1      localhost
#       ::1            localhost
10.82.129.120 rms.'...' .com # modified by IrmTool at 2015-01-18 17:55:13
```

For details about the solution, see [Step 4](#).

## Solution

**Step 1** The IP address for connecting to ZooKeeper is incorrectly configured.

**Step 2** Change the IP address for connecting to ZooKeeper in the Consumer configuration and make it consistent with MRS configuration.

- Flume  
server.sources.Source02.type=org.apache.flume.source.kafka.KafkaSource  
server.sources.Source02.zookeeperConnect=192.168.234.231:2181/kafka  
server.sources.Source02.topic = test  
server.sources.Source02.groupId = test\_01
- Spark  
val zkQuorum = "192.168.234.231:2181/kafka"
- Storm  
BrokerHosts brokerHosts = new ZKHosts("192.168.234.231:2181/kafka");
- Consumer API  
zookeeper.connect="192.168.234.231:2181/kafka"

**Step 3** Rebalance is abnormal.

Multiple Consumers in the same consumer group are successively started and consume data of multiple partitions at the same time, load balancing is performed for Consumers when consumers are fewer than partitions.

The temporary node where the Consumer is stored on ZooKeeper determines read/write permission of which partition of which topic the Consumer has. The path is `/consumers/consumer-group-xxx/owners/topic-xxx/x`.

After the load balancing is triggered, the original Consumer will be recalculated and release occupied partitions, which takes a while. Therefore, new Consumers may fail to preempt the partitions.

**Table 16-1** Parameters

Name	Function	Default Value
rebalance.max.retries	Maximum number of rebalance retries	4
rebalance.backoff.ms	Interval for each rebalance retry	2000
zookeeper.session.timeout.ms	Maximum time allowed to create a session with ZooKeeper	15000

Set the preceding parameters to higher values. The following is for your reference:

```
zookeeper.session.timeout.ms = 45000
rebalance.max.retries = 10
rebalance.backoff.ms = 5000
```

Parameter setting must comply with the following rule:

**rebalance.max.retries \* rebalance.backoff.ms > zookeeper.session.timeout.ms**

**Step 4** The network is abnormal.

In the **hosts** file, mapping between the hostname and IP address is not configured. As a result, information cannot be obtained when using the hostname for access.

**Step 5** Add the hostname to the **hosts** file and make it correspond to the IP address.

- Linux

```
127.0.0.1    localhost

192.168.0.131 192-168-0-131
192.168.0.51  192-168-0-51
192.168.0.122 192-168-0-122
192.168.234.231 192-168-234-231
```

- Windows

```
# For example:
#
# 192.168.94.97 rhino.acme.com # source server
# 192.168.63.10 x.acme.com # x client host

# localhost name resolution is handled within DNS itself.
# 127.0.0.1 localhost
# ::1 localhost
10.82.129.120 rms [REDACTED].com # modified by IrmTool at 2015-01-18 17:55:13
192.168.234.231 192-168-234-231
```

----End

## 16.8 Consumer Fails to Consume Data and Remains in the Waiting State

### Symptom

An MRS cluster is installed, and ZooKeeper and Kafka are installed in the cluster.

When the Consumer consumes data from Kafka, the client stays in the Waiting state.

### Possible Causes

1. The Kafka service is abnormal.
2. The Consumer of the client adopts non-security access and access is disabled on the server.
3. The Consumer of the client adopts non-security access and ACL is set for Kafka topics.

### Cause Analysis

The possible reasons why the Consumer fails to consume data from Kafka may be related to the Consumer or Kafka.

1. Check the Kafka service status:
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
2. Check the Consumer client log. It is found that the information about the frequent connections and disconnections to the Broker node is printed, as shown in the following output.

```
[root@10-10-144-2 client]# kafka-console-consumer.sh --topic test --zookeeper 10.5.144.2:2181/kafka
--from-beginning
[2017-03-07 09:22:00,658] INFO Fetching metadata from broker BrokerEndPoint(1,10.5.144.2,9092)
with correlation id 26 for 1 topic(s) Set(test) (kafka.client.ClientUtils$)
[2017-03-07 09:22:00,659] INFO Connected to 10.5.144.2:9092 for producing
```



```
(kafka.producer.SyncProducer)
[2017-03-07 09:22:00,659] INFO Disconnecting from 10.5.144.2:9092 (kafka.producer.SyncProducer)
```

Consumer accesses Kafka using port 9092, which is a non-security port.

3. On Manager, check the current Kafka cluster configuration. It is found that the customized configuration **allow.everyone.if.no.acl.found=false** is not configured.
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka > Configurations**.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka > Configurations**.

4. If ACL is set to **false**, port 9092 cannot be used for access.

5. Check the Consumer client log. It is found that the information about the frequent connections and disconnections to the Broker node is printed, as shown in the following output.

```
[root@10-10-144-2 client]# kafka-console-consumer.sh --topic test_acl --zookeeper 10.5.144.2:2181/
kafka --from-beginning
[2017-03-07 09:49:16,992] INFO Fetching metadata from broker BrokerEndPoint(2,10.5.144.3,9092)
with correlation id 16 for 1 topic(s) Set(topic_acl) (kafka.client.ClientUtils$)
[2017-03-07 09:49:16,993] INFO Connected to 10.5.144.3:9092 for producing
(kafka.producer.SyncProducer)
[2017-03-07 09:49:16,994] INFO Disconnecting from 10.5.144.3:9092 (kafka.producer.SyncProducer)
```

The Consumer accesses Kafka using port 21005, which is a non-security port.

6. Run the client command to check the ACL permission of the topic.

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/
kafka --list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
User:test_user has Allow permission for operations: Write from hosts: *
```

If ACL is set for the topic, port 9092 cannot be used for access.

7. The following information is printed in the Consumer client log:

```
[root@10-10-144-2 client]# kafka-console-consumer.sh --topic topic_acl --bootstrap-server
10.5.144.2:21007 --consumer.config /opt/client/Kafka/kafka/config/consumer.properties --from-
beginning --new-consumer
[2017-03-07 10:19:18,478] INFO Kafka version : 0.9.0.0 (org.apache.kafka.common.utils.AppInfoParser)
[2017-03-07 10:19:18,478] INFO Kafka commitId : unknown
(org.apache.kafka.common.utils.AppInfoParser)
```

The Consumer uses port 21007 to access Kafka.

8. Run the client command **klist** to query the current authenticated user.

```
[root@10-10-144-2 client]# klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: test@HADOOP.COM

Valid starting Expires Service principal
01/25/17 11:06:48 01/26/17 11:06:45 krbtgt/HADOOP.COM@HADOOP.COM
```

The **test** user is used in this example.

9. Run the client command to check the ACL permission of the topic.

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:24002/
kafka --list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
User:test_user has Allow permission for operations: Write from hosts: *
User:ttest_user has Allow permission for operations: Read from hosts: *
```

If ACL is set for the topic, user **test** does not have the permission to perform the Consumer operation.

For details about the solution, see [Step 2](#).

10. Log in to Kafka Broker using SSH.

Run the `cd /var/log/Bigdata/kafka/broker` command to go to the log directory.

Check the `kafka-authorizer.log` file. It shows that the user does not belong to the `kafka` or `kafkaadmin` group.

```
2017-01-25 13:26:33,648 | INFO | [kafka-request-handler-0] | The principal is test, belongs to Group : [hadoop, ficommon] | kafka.authorizer.logger (SimpleAclAuthorizer.scala:169)
2017-01-25 13:26:33,648 | WARN | [kafka-request-handler-0] | The user is not belongs to kafka or kafkaadmin group, authorize failed! | kafka.authorizer.logger (SimpleAclAuthorizer.scala:170)
```

For details about the solution, see [Step 3](#).

## Solution

**Step 1** Set `allow.everyone.if.no.acl.found` to `true` and restart the Kafka service.

**Step 2** Use the account with permission for login.

Example:

```
kinit test_user
```

Alternatively, grant the user with related permission.

---

### NOTICE

This operation must be performed by the Kafka administrator (belonging to the `kafkaadmin` group).

Example:

```
kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/kafka --topic topic_acl --consumer --add --allow-principal User:test --group test
```

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=8.5.144.2:2181/kafka --list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
User:test_user has Allow permission for operations: Write from hosts: *
User:test has Allow permission for operations: Describe from hosts: *
User:test has Allow permission for operations: Write from hosts: *
User:test has Allow permission for operations: Read from hosts: *
```

**Step 3** Add the user to the `kafka` or `kafkaadmin` group.

----End

## 16.9 SparkStreaming Fails to Consume Kafka Messages, and "Error getting partition metadata" Is Displayed

### Symptom

When SparkStreaming is used to consume messages of a specified topic in Kafka, data cannot be obtained from Kafka. The message "Error getting partition metadata" is displayed.

```
Exception in thread "main" org.apache.spark.SparkException: Error getting partition metadata for 'testtopic'. Does the topic exist?
org.apache.spark.streaming.kafka.KafkaCluster$$anonfun$checkErrors$1.apply(KafkaCluster.scala:366)
org.apache.spark.streaming.kafka.KafkaCluster$$anonfun$checkErrors$1.apply(KafkaCluster.scala:366)
scala.util.Either.fold(Either.scala:97)
org.apache.spark.streaming.kafka.KafkaCluster$.checkErrors(KafkaCluster.scala:365)
org.apache.spark.streaming.kafka.KafkaUtils$.createDirectStream(KafkaUtils.scala:422)
com.xxx.bigdata.spark.examples.FemaleInfoCollectionPrint$.main(FemaleInfoCollectionPrint.scala:45)
com.xxx.bigdata.spark.examples.FemaleInfoCollectionPrint.main(FemaleInfoCollectionPrint.scala)
sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
java.lang.reflect.Method.invoke(Method.java:498)
org.apache.spark.deploy.SparkSubmit$.org$apache$spark$deploy$SparkSubmit$
$runMain(SparkSubmit.scala:762)
org.apache.spark.deploy.SparkSubmit$.doRunMain$1(SparkSubmit.scala:183)
org.apache.spark.deploy.SparkSubmit$.submit(SparkSubmit.scala:208)
org.apache.spark.deploy.SparkSubmit$.main(SparkSubmit.scala:123)
org.apache.spark.deploy.SparkSubmit.main(SparkSubmit.scala)
```

## Possible Causes

1. The Kafka service is abnormal.
2. The Consumer client adopts non-security access and access is disabled on the server.
3. The Consumer of the client adopts non-security access and ACL is set for Kafka topics.

## Cause Analysis

1. Check the Kafka service status:
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
2. On Manager, check the current Kafka cluster configuration. It is found that **allow.everyone.if.no.acl.found** is not configured or is set to **false**.
3. If it is set to **false**, the Kafka non-secure port 21005 cannot be used for access.
4. Run the client command to check the ACL permission of the topic.

```
[root@10-10-144-2 client]# kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/
kafka --list --topic topic_acl
Current ACLs for resource `Topic:topic_acl`:
User:test_user has Allow permission for operations: Describe from hosts: *
User:test_user has Allow permission for operations: Write from hosts: *
```

If an ACL is configured for a topic, the Kafka non-secure port 21005 cannot be used to access the topic.

## Solution

**Step 1** Add the customized configuration **allow.everyone.if.no.acl.found** or change its value to **true** and restart the Kafka service.

**Step 2** Delete the ACL configured for the topic.

Example:

### kinit test\_user

#### NOTICE

This operation must be performed by the Kafka administrator (belonging to the **kafkaadmin** group).

Example:

```
kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/kafka  
--remove --allow-principal User:test_user --producer --topic topic_acl
```

```
kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/kafka  
--remove --allow-principal User:test_user --consumer --topic topic_acl --group  
test
```

----End

## 16.10 Consumer Fails to Consume Data in a Newly Created Cluster, and Message "GROUP\_COORDINATOR\_NOT\_AVAILABLE" Is Displayed

### Symptom

A Kafka cluster is created, and two Broker nodes are deployed. The Kafka client can be used for production but cannot be used for consumption. The Consumer fails to consume data, and the message "GROUP\_COORDINATOR\_NOT\_AVAILABLE" is displayed. The key log is as follows:

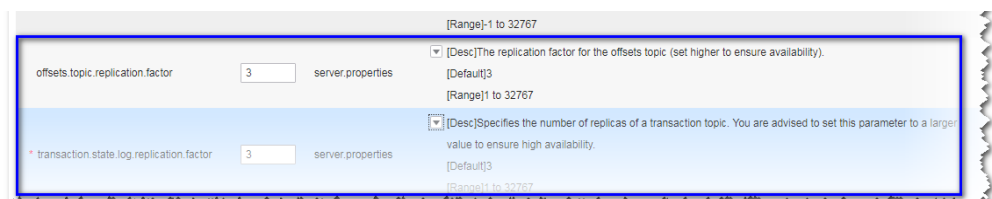
```
2018-05-12 10:58:42,561 | INFO | [kafka-request-handler-3] | [GroupCoordinator 2]: Preparing to restabilize  
group DemoConsumer with old generation 118 | kafka.coordinator.GroupCoordinator (Logging.scala:68)  
2018-05-12 10:59:13,562 | INFO | [executor-Heartbeat] | [GroupCoordinator 2]: Preparing to restabilize  
group DemoConsumer with old generation 119 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
```

### Possible Causes

The **\_\_consumer\_offsets** cannot be created.

### Cause Analysis

1. As indicated by the log, a large number of **\_\_consumer\_offset** creation operations failed.
2. The number of Brokers for the cluster is 2.
3. However, the number of replicas for the **\_\_consumer\_offset** topic is 3. Therefore, the creation fails.



## Solution

Expand the cluster to at least three streaming core nodes or perform the following steps to modify service configuration parameters:

**Step 1** Go to the service configuration page.

- MRS Manager: Log in to MRS Manager, choose **Services > Kafka > Service Configuration**, and select **All** from **Type**.
- FusionInsight Manager: Log in to FusionInsight Manager. Choose **Cluster > Services > Kafka**. Click **Configurations** and select **All Configurations**.

**Step 2** Search for **offsets.topic.replication.factor** and **transaction.state.log.replication.factor** and change their values to **2**.

**Step 3** Save the configuration and restart the service or instance whose configuration has expired.

----End

## 16.11 SparkStreaming Fails to Consume Kafka Messages, and Message "Couldn't find leader offsets" Is Displayed

### Symptom

When SparkStreaming is used to consume messages of a specified topic in Kafka, data cannot be obtained from Kafka.

The following error message is displayed: Couldn't find leader offsets.

```
2018-05-30 12:01:17,816 | INFO | [Driver] | Reconnect due to socket error: java.net.SocketTimeoutException | kafka.utils.Logging$class.info(Logging.scala:68)
2018-05-30 12:01:47,859 | ERROR | [Driver] | User class threw exception: org.apache.spark.SparkException: java.net.SocketTimeoutException
org.apache.spark.SparkException: Couldn't find leader offsets for Set([STEB, 57], [STEB, 21]) | org.apache.spark.Logging$class.logError(Logging.scala:96)
org.apache.spark.SparkException: java.net.SocketTimeoutException
org.apache.spark.SparkException: Couldn't find leader offsets for Set([STEB, 57], [STEB, 21])
at org.apache.spark.streaming.kafka.KafkaCluster$$anonfun$checkErrors$1.apply(KafkaCluster.scala:366)
at org.apache.spark.streaming.kafka.KafkaCluster$$anonfun$checkErrors$1.apply(KafkaCluster.scala:366)
at scala.util.Either.fold(Either.scala:97)
at org.apache.spark.streaming.kafka.KafkaCluster$.checkErrors(KafkaCluster.scala:365)
at org.apache.spark.streaming.kafka.KafkaUtils$.createDirectStream(KafkaUtils.scala:422)
at org.apache.spark.streaming.kafka.KafkaUtils$.createDirectStream(KafkaUtils.scala:532)
at org.apache.spark.streaming.kafka.KafkaUtils.createDirectStream(KafkaUtils.scala)
at com.stk.bigdata.sparkstreaming.notify.SparkAlarmControlwz.main(SparkAlarmControlwz.java:194)
at com.stk.bigdata.sparkstreaming.submit.SparkNotifyA.main(SparkNotifyA.java:14)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at org.apache.spark.deploy.yarn.ApplicationMaster$$anon$2.run(ApplicationMaster.scala:540)
2018-05-30 12:01:47,863 | INFO | [Driver] | Final app status: FAILED, exitCode: 15, (reason: User class threw exception: org.apache.spark.SparkException: java:
org.apache.spark.SparkException: Couldn't find leader offsets for Set([STEB, 57], [STEB, 21])) | org.apache.spark.Logging$class.logInfo(Logging.scala:59)
2018-05-30 12:01:47,866 | INFO | [pool-1-thread-1] | Invoking stop() from shutdown hook | org.apache.spark.Logging$class.logInfo(Logging.scala:59)
```

### Possible Causes

- The Kafka service is abnormal.
- The network is abnormal.
- The Kafka topic is abnormal.

### Cause Analysis

**Step 1** On Manager, check the status of the Kafka cluster. The status is **Good**, and the monitoring metrics are correctly displayed.

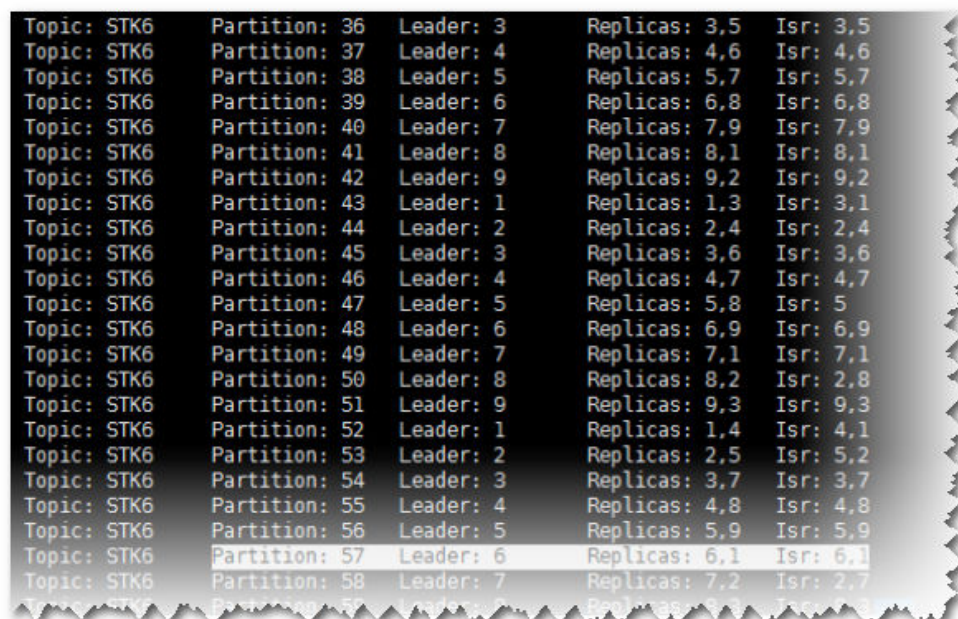
**Step 2** View the error topic information in the SparkStreaming log.

Run the Kafka commands to obtain the topic assignment information and copy synchronization information, and check the return result.

**kafka-topics.sh --describe --zookeeper <zookeeper> --topic <topic name>**

If information in the following figure is displayed, the topic is normal. All partitions have normal leader information.

**Figure 16-10** Topic distribution information and copy synchronization information



**Step 3** Check whether the network connection between the client and Kafka cluster is normal. If no, contact the network team to rectify the fault.

**Step 4** Log in to Kafka Broker using SSH.

Run the **cd /var/log/Bigdata/kafka/broker** command to go to the log directory.

Check on **server.log** indicates that the error message is displayed in the log shown in the following figure.

```
2018-05-30 12:02:00,246 | ERROR | [kafka-network-thread-6-PLAINTEXT-3] | Processor got uncaught exception. | kafka.network.Processor (Logging.scala:103)
java.lang.OutOfMemoryError: Direct buffer memory
at java.nio.Bits.reserveMemory(Bits.java:694)
at java.nio.DirectByteBuffer.<init>(DirectByteBuffer.java:123)
at java.nio.ByteBuffer.allocateDirect(ByteBuffer.java:311)
at sun.nio.ch.Util.getTemporaryDirectBuffer(Util.java:241)
at sun.nio.ch.IOUtil.read(IOUtil.java:195)
at sun.nio.ch.SocketChannelImpl.read(SocketChannelImpl.java:380)
```

```
at
org.apache.kafka.common.network.PlaintextTransportLayer.read(PlaintextTransport
Layer.java:110)
```

**Step 5** On Manager, check the configuration of the current Kafka cluster.

- MRS Manager: Log in to MRS Manager and choose **Services > Kafka > Service Configuration**. Set **Type** to **All**. The value of **-XX:MaxDirectMemorySize** in **KAFKA\_JVM\_PERFORMANCE\_OPTS** is **1G**.
- FusionInsight Manager: Log in to FusionInsight Manager. Choose **Cluster > Services > Kafka > Configurations > All Configurations**. The value of **-XX:MaxDirectMemorySize** in **KAFKA\_JVM\_PERFORMANCE\_OPTS** is **1G**.

**Step 6** If the direct memory is too small, an error is reported. Once the direct memory overflows, the node cannot process new requests. As a result, other nodes or clients fail to access the node due to timeout.

----End

## Solution

**Step 1** Log in to Manager and go to the Kafka configuration page.

**Step 2** Set **Type** to **All**, and search for and change the value of **KAFKA\_JVM\_PERFORMANCE\_OPTS**.

**Step 3** Save the configuration and restart the service or instance whose configuration has expired.

----End

## 16.12 Consumer Fails to Consume Data and Message "SchemaException: Error reading field" Is Displayed

### Symptom

When a Consumer consumes messages of a specified topic in Kafka, the Consumer cannot obtain data from Kafka.

The following error message is displayed:

```
org.apache.kafka.common.protocol.types.SchemaException: Error reading field 'brokers': Error reading field 'host': Error reading string of length 28271, only 593 bytes available
```

```
Exception in thread "Thread-0" org.apache.kafka.common.protocol.types.SchemaException: Error reading field 'brokers': Error reading field 'host': Error reading string of length 28271, only 593 bytes available
at org.apache.kafka.common.protocol.types.Schema.read(Schema.java:73)
at org.apache.kafka.clients.NetworkClient.parseResponse(NetworkClient.java:380)
at org.apache.kafka.clients.NetworkClient.handleCompletedReceives(NetworkClient.java:449)
at org.apache.kafka.clients.NetworkClient.poll(NetworkClient.java:269)
at
org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.clientPoll(ConsumerNetworkClient.java:360)
at
org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.poll(ConsumerNetworkClient.java:224)
at
org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.poll(ConsumerNetworkClient.java:192)
at
org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.poll(ConsumerNetworkClient.java:163)
at org.apache.kafka.clients.consumer.internals.AbstractCoordinator.ensureCoordinatorReady(AbstractCoordinator.java:179)
at org.apache.kafka.clients.consumer.KafkaConsumer.pollOnce(KafkaConsumer.java:973)
at org.apache.kafka.clients.consumer.KafkaConsumer.poll(KafkaConsumer.java:937)
at KafkaNew.Consumer$ConsumerThread.run(Consumer.java:40)
```

## Possible Causes

The JAR versions of the client and server are inconsistent.

## Solution

Modify the Kafka JAR package in the Consumer application to ensure that it is the same as that on the server.

# 16.13 Kafka Consumer Loses Consumed Data

## Symptom

A user saves the consumed data to the database and finds that the data is inconsistent with the production data. Therefore, it is suspected that some of Kafka's consumed data is lost.

## Possible Causes

- The service code is incorrect.
- An exception occurs when Kafka production data is written.
- The Kafka consumption data is abnormal.

## Solution

Check Kafka.

- Step 1** Observe the changes of the written and consumed offset through **consumer-groups.sh**. (Produce a certain number of messages, and consume these messages on the client to observe the changes of the offset.)

```
2019-04-08 14:29:28,341 WARN [Principal:null]: TGT renewal thread has been interrupted and will exit. (org.apache.kafka.common.security.kerberos.KerberosLogin)
root@mslqdata033 kafka: /bin/kafka-consumer-groups.sh --describe --bootstrap-server 10.3.1.49:21007 --new-consumer --group ghdshoj --command-config config/consum
Properties
Note: This will only show information about consumers that use the Java consumer API (non-ZooKeeper-based consumers).

TOPIC          PARTITION  CURRENT-OFFSET  LOG-END-OFFSET  LAG             CONSUMER-ID                                     HOST
KMT-ID
LMSJDSDB      0          290078541       290078541       0               consumer-1-7bb54edf-9cbb-4d58-989b-1b4e6607217e /10.2.1.180
sumer-1
LMSJDSDB      1          281608671       281608671       0               consumer-1-7bb54edf-9cbb-4d58-989b-1b4e6607217e /10.2.1.180
sumer-1
LMSJDSDB      2          283880519       283880519       0               consumer-1-7bb54edf-9cbb-4d58-989b-1b4e6607217e /10.2.1.180
sumer-1
```

- Step 2** Create a consumption group, use the client to consume messages, and view the consumed messages.

new-consumer:

```
kafka-console-consumer.sh --topic <topic name> --bootstrap-server <IP1:PORT, IP2:PORT,...> --new-consumer --consumer.config <config file>
```

----End

Check the service code.

- Step 1** Check whether an error is reported when the offset is submitted on the client.
- Step 2** If no error is reported, add a printing message to the API that is consumed, and print only the key to view the lost data.

----End



## 16.14 Failed to Start Kafka Due to Account Lockout

### Symptom

The Kafka service fails to be started in the newly created MRS cluster.

The service startup log shows that the authentication fails.

```
/home/omm/kerberos/bin/kinit -k -t ${BIGDATA_HOME}/etc/2_15_ Broker /kafka.keytab kafka/  
hadoop.hadoop.com -c ${BIGDATA_HOME}/etc/2_15_ Broker /11846 failed.  
export key tab file for kafka/hadoop.hadoop.com failed.export and check keytab file failed, errMsg=}}] for  
Broker #192.168.1.92@192-168-1-92.  
[2015-07-11 02:34:33] RoleInstance started failure for ROLE[name: Broker].  
[2015-07-11 02:34:34] Failed to complete the instances start operation. Current operation entities: [Broker  
#192.168.1.92@192-168-1-92], Failure entites : [Broker #192.168.1.92@192-168-1-92].Operation  
Failed.Failed to complete the instances start operation. Current operation entities:  
[Broker#192.168.1.92@192-168-1-92], Failure entites: [Broker #192.168.1.92@192-168-1-92].
```

### Cause Analysis

The Kerberos log `/var/log/Bigdata/kerberos/krb5kdc.log` shows that IP addresses outside the cluster set up connections using a Kafka account, resulting in consecutive authentication failures and account lockout.

```
Jul 11 02:49:16 192-168-1-91 krb5kdc[1863](info): AS_REQ (2 etypes {18 17}) 192.168.1.93:  
NEEDED_PREAUTH: kafka/hadoop.hadoop.com@HADOOP.COM for krbtgt/HADOOP.COM@HADOOP.COM,  
Additional pre-authentication required  
Jul 11 02:49:16 192-168-1-91 krb5kdc[1863](info): preauth (encrypted_timestamp) verify failure: Decrypt  
integrity check failed  
Jul 11 02:49:16 192-168-1-91 krb5kdc[1863](info): AS_REQ (2 etypes {18 17}) 192.168.1.93:  
PREAUTH_FAILED: kafka/hadoop.hadoop.com@HADOOP.COM for krbtgt/HADOOP.COM@HADOOP.COM,  
Decrypt integrity check failed
```

### Solution

- Step 1** Check the IP address of the node that connects to Kafka outside the cluster, for example, 192.168.1.93 in the example.
- Step 2** Log in to the node outside the cluster and disable Kafka authentication on the node.
- Step 3** Wait 5 minutes for the account to be unlocked.
- Step 4** Restart the Kafka service.

----End

## 16.15 Kafka Broker Reports Abnormal Processes and the Log Shows "IllegalArgumentExpection"

### Symptom

The Process Fault alarm is reported on Manager. Check whether the faulty process is Kafka Broker.

## Possible Causes

Broker configuration is abnormal.

## Cause Analysis

1. On Manager, obtain the host information on the alarm page.
2. Log in to Kafka Broker using SSH. Run the `cd /var/log/Bigdata/kafka/broker` command to go to the log directory.

Check the **server.log** file. It is found that the "IllegalArgumentException" exception is thrown in the following log stating "java.lang.IllegalArgumentException: requirement failed: replica.fetch.max.bytes should be equal or greater than message.max.bytes."

```
2017-01-25 09:09:14,930 | FATAL | [main] | | kafka.Kafka$ (Logging.scala:113)
java.lang.IllegalArgumentException: requirement failed: replica.fetch.max.bytes should be equal or
greater than message.max.bytes
    at scala.Predef$.require(Predef.scala:233)
    at kafka.server.KafkaConfig.validateValues(KafkaConfig.scala:959)
    at kafka.server.KafkaConfig.<init>(KafkaConfig.scala:944)
    at kafka.server.KafkaConfig$.fromProps(KafkaConfig.scala:701)
    at kafka.server.KafkaConfig$.fromProps(KafkaConfig.scala:698)
    at kafka.server.KafkaServerStartable$.fromProps(KafkaServerStartable.scala:28)
    at kafka.Kafka$.main(Kafka.scala:60)
    at kafka.Kafka.main(Kafka.scala)
```

Kafka requires that **replica.fetch.max.bytes** be greater than or equal to **message.max.bytes**.

3. On the Kafka configuration page, select **All Configurations**. All Kafka configurations are displayed. Search for **message.max.bytes** and **replica.fetch.max.bytes**. It is found that the value of **replica.fetch.max.bytes** is less than that of **message.max.bytes**.

## Solution

**Step 1** Go to the Kafka configuration page.

- For versions earlier than MRS 3.x: Log in to MRS Manager and choose **Services > Kafka > Service Configuration > All Configurations**.
- For MRS 3.x or later: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka > Configurations > All Configurations**.

**Step 2** Search for and modify the **replica.fetch.max.bytes** parameter to ensure that its value is greater than or equal to that of **message.max.bytes**. In this way, replicas of partitions on different brokers can be synchronized to all messages.

**Step 3** Save the configuration and check whether there is any service whose configuration has expired in the cluster. If yes, restart the corresponding service or role instance for the configuration to take effect.

**Step 4** Modify **fetch.message.max.bytes** in the Consumer service application to ensure that the value of **fetch.message.max.bytes** is greater than or equal to that of **message.max.bytes**.

----End

## 16.16 Kafka Topics Cannot Be Deleted

### Symptom

When running the following command on the Kafka client to delete topics, it is found that the topics cannot be deleted.

```
kafka-topics.sh --delete --topic test --zookeeper 192.168.234.231:2181/kafka
```

### Possible Causes

- The command for connecting the client to ZooKeeper is incorrect.
- Kafka is abnormal and some Kafka nodes are stopped.
- Perform the following operations when Kafka server configurations cannot be deleted.
- Perform the following operations when Kafka configurations are automatically created and the Producer is not stopped.

### Cause Analysis

1. After the client command is run, the "ZkTimeoutException" exception is reported.

```
[2016-03-09 10:41:45,773] WARN Can not get the principle name from server 192.168.234.231  
(org.apache.zookeeper.ClientCnxn)  
Exception in thread "main" org.I0ltec.zkclient.exception.ZkTimeoutException: Unable to connect to  
zookeeper server within timeout: 30000  
at org.I0ltec.zkclient.ZkClient.connect(ZkClient.java:880)  
at org.I0ltec.zkclient.ZkClient.<init>(ZkClient.java:98)  
at org.I0ltec.zkclient.ZkClient.<init>(ZkClient.java:84)  
at kafka.admin.TopicCommand$.main(TopicCommand.scala:51)  
at kafka.admin.TopicCommand.main(TopicCommand.scala)
```

For details about the solution, see [Step 1](#).

2. Run the following query command on the client:

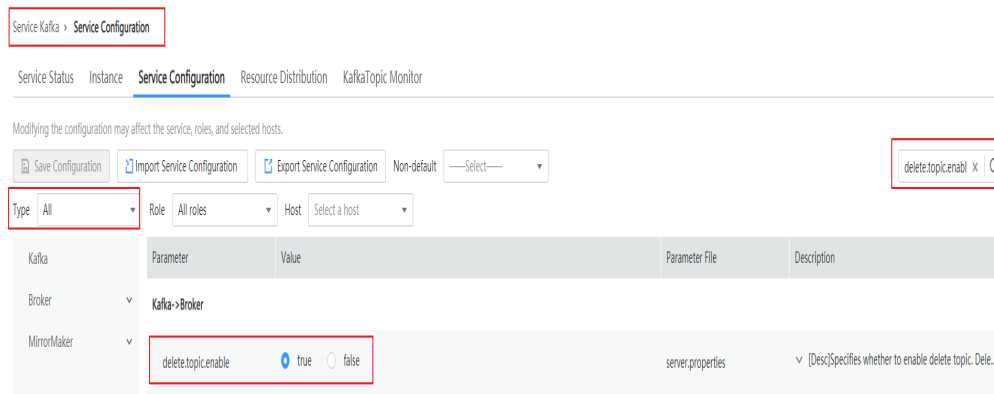
```
kafka-topics.sh --list --zookeeper 192.168.0.122:2181/kafka  
test - marked for deletion
```

On Manager, check the running status of Kafka Broker instances.

Run the `cd /var/log/Bigdata/kafka/broker` command to go to the log directory of node **RunningAsController**. Locate **ineligible for deletion: test** in the **controller.log** file.

```
2016-03-09 11:11:26,228 | INFO | [main] | [Controller 1]: List of topics to be deleted: |  
kafka.controller.KafkaController (Logging.scala:68)  
2016-03-09 11:11:26,229 | INFO | [main] | [Controller 1]: List of topics ineligible for deletion: test |  
kafka.controller.KafkaController (Logging.scala:68)
```

3. On Manager, view the **delete.topic.enable** status of Broker.



For details about the solution, see [Step 2](#).

- Run the following query command on the client:

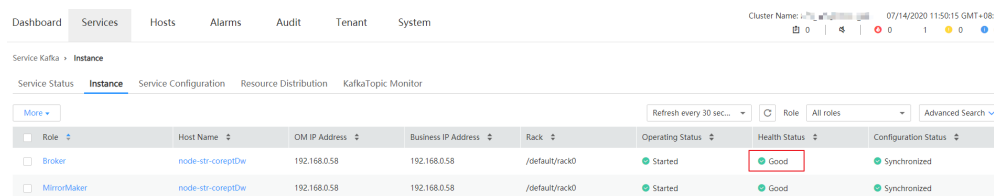
```
kafka-topics.sh --describe -topic test --zookeeper 192.168.0.122:2181/kafka
```

```
Topic:test      PartitionCount:10  ReplicationFactor:2  Configs:
Topic: test    Partition: 0      Leader: -1           Replicas: 1,2  Isr:
Topic: test    Partition: 1      Leader: -1           Replicas: 2,3  Isr:
Topic: test    Partition: 2      Leader: -1           Replicas: 3,1  Isr:
Topic: test    Partition: 3      Leader: -1           Replicas: 1,3  Isr:
Topic: test    Partition: 4      Leader: -1           Replicas: 2,1  Isr:
Topic: test    Partition: 5      Leader: -1           Replicas: 3,2  Isr:
Topic: test    Partition: 6      Leader: -1           Replicas: 1,2  Isr:
Topic: test    Partition: 7      Leader: -1           Replicas: 2,3  Isr:
Topic: test    Partition: 8      Leader: -1           Replicas: 3,1  Isr:
Topic: test    Partition: 9      Leader: -1           Replicas: 1,3  Isr:
```

Go to the log directory of node **RunningAsController**. Locate **marked ineligible for deletion** in the **controller.log** file.

```
2016-03-10 11:11:17,989 | INFO | [delete-topics-thread-3] | [delete-topics-thread-3], Handling
deletion for topics test | kafka.controller.TopicDeletionManager$DeleteTopicsThread (Logging.scala:68)
2016-03-10 11:11:17,990 | INFO | [delete-topics-thread-3] | [delete-topics-thread-3], Not retrying
deletion of topic test at this time since it is marked ineligible for deletion |
kafka.controller.TopicDeletionManager$DeleteTopicsThread (Logging.scala:68)
```

- On Manager, query the Broker status.



It can be seen that a Broker is in the Stopped state. In this case, delete the topic and ensure that Brokers where partitions of the topic reside must be in the Good state.

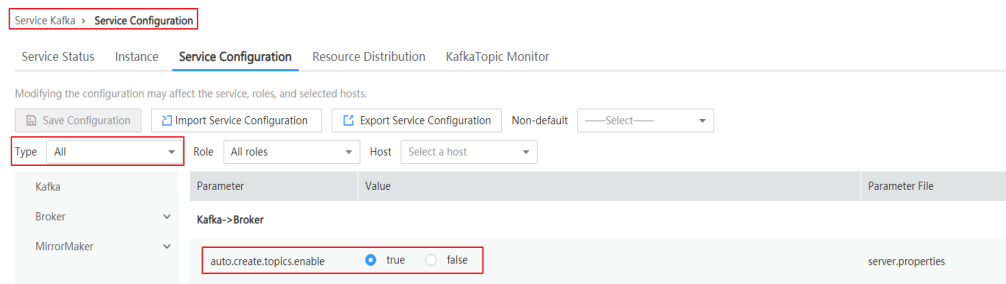
For details about the solution, see [Step 3](#).

- Go to the log directory of node **RunningAsController**. Locate **Deletion successfully** in the **controller.log** file. If **New topics:[Set(test)]** is displayed again, it indicates that the topic is created again.

```
2016-03-10 11:33:35,208 | INFO | [delete-topics-thread-3] | [delete-topics-thread-3], Deletion of topic
test successfully completed | kafka.controller.TopicDeletionManager$DeleteTopicsThread
(Logging.scala:68)
```

```
2016-03-10 11:33:38,501 | INFO | [ZkClient-
EventThread-19-192.168.0.122:2181,160.172.0.52:2181,160.172.0.51:2181/kafka] |
[TopicChangeListener on Controller 3]: New topics: [Set(test)], deleted topics: [Set()], new partition
replica assignment
```

- Use Manager to query the topic creation configuration of Broker.



It is confirmed that the application that performs operations on the topic is not stopped.

For details about the solution, see [Step 4](#).

## Solution

**Step 1** Perform the following operations when connection to ZooKeeper fails.

When the connection between the Kafka client and ZooKeeper times out, run the ping command to check whether the Kafka client can connect to ZooKeeper. Check the network connection between the client and ZooKeeper.

If the network connection fails, check the ZooKeeper service information on Manager.

**Figure 16-11** ZooKeeper service information

The screenshot shows the 'Instance' page for ZooKeeper. At the top, there are tabs for 'Service Status', 'Instance', 'Service Configuration', and 'Resource Distribution'. Below the tabs, there is a 'More' dropdown and a 'Refresh every 30 sec...' dropdown. The table shows the following instance information:

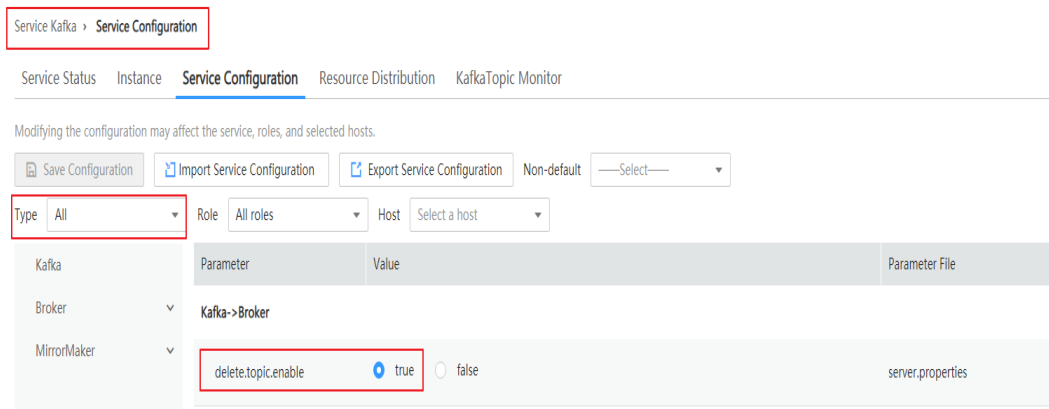
Role	Host Name	OM IP Address	Business IP Address	Rack	Operating Status
<input type="checkbox"/> quorumpeer	node-master1Qb5V	192.168.0.210	192.168.0.210	/default/rackca2e	<span style="color: green;">●</span> Started

If ZooKeeper is improperly configured, change the ZooKeeper IP address in the client command.

**Step 2** Perform the following operations when Kafka server configurations cannot be deleted.

On Manager, change the value of **delete.topic.enable** to **true**. Save the configurations and restart the service.

**Figure 16-12** Modifying delete.topic.enable.



The client query command does not contain **Topic:test**.

```
kafka-topics.sh --list --zookeeper 192.168.0.122:24002/kafka
```

Go to the log directory of node **RunningAsController**. Locate **Deletion of topic test successfully** in the **controller.log** file.

```
2016-03-10 10:39:40,665 | INFO | [delete-topics-thread-3] | [Partition state machine on Controller 3]:
Invoking state change to OfflinePartition for partitions [test,2],[test,15],[test,6],[test,16],[test,12],[test,7],
[test,10],[test,13],[test,9],[test,19],[test,3],[test,5],[test,1],[test,0],[test,17],[test,8],[test,4],[test,11],[test,14],
[test,18] | kafka.controller.PartitionStateMachine (Logging.scala:68)
2016-03-10 10:39:40,668 | INFO | [delete-topics-thread-3] | [Partition state machine on Controller 3]:
Invoking state change to NonExistentPartition for partitions [test,2],[test,15],[test,6],[test,16],[test,12],
[test,7],[test,10],[test,13],[test,9],[test,19],[test,3],[test,5],[test,1],[test,0],[test,17],[test,8],[test,4],[test,11],
[test,14],[test,18] | kafka.controller.PartitionStateMachine (Logging.scala:68)
2016-03-10 10:39:40,977 | INFO | [delete-topics-thread-3] | [delete-topics-thread-3], Deletion of topic test
successfully completed | kafka.controller.TopicDeletionManager$DeleteTopicsThread (Logging.scala:68)
```

**Step 3** Some Kafka nodes are stopped or faulty.

Start the stopped Broker instances.

The client query command does not contain **Topic:test**.

```
kafka-topics.sh --list --zookeeper 192.168.0.122:24002/kafka
```

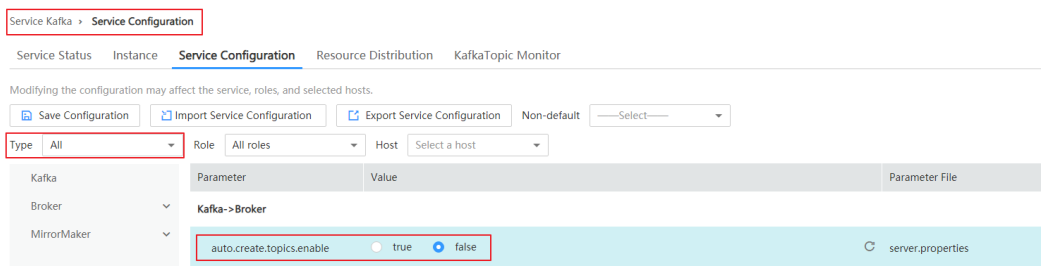
Go to the log directory of node **RunningAsController**. Locate **Deletion of topic test successfully** in the **controller.log** file.

```
2016-03-10 11:17:56,463 | INFO | [delete-topics-thread-3] | [Partition state machine on Controller 3]:
Invoking state change to NonExistentPartition for partitions [test,4],[test,1],[test,8],[test,2],[test,5],[test,9],
[test,7],[test,6],[test,0],[test,3] | kafka.controller.PartitionStateMachine (Logging.scala:68)
2016-03-10 11:17:56,726 | INFO | [delete-topics-thread-3] | [delete-topics-thread-3], Deletion of topic test
successfully completed | kafka.controller.TopicDeletionManager$DeleteTopicsThread (Logging.scala:68)
```

**Step 4** Perform the following operations when Kafka configurations are automatically created and the Producer is not stopped.

Stop related applications, change the value of **auto.create.topics.enable** to **false** on Manager. Save the configuration and restart the service.

**Figure 16-13** Modifying auto.create.topics.enable



**Step 5** Perform the delete operation again.

----End

## 16.17 Error "AdminOperationException" Is Displayed When a Kafka Topic Is Deleted

### Symptom

When running the following command on the Kafka client to set the ACL for a topic, it is found that the ACL cannot be set.

```
kafka-topics.sh --delete --topic test4 --zookeeper 10.5.144.2:2181/kafka
```

The error message "ERROR kafka.admin.AdminOperationException: Error while deleting topic test4" is displayed.

Details are as follows:

```
Error while executing topic command : Error while deleting topic test4
[2017-01-25 14:00:20,750] ERROR kafka.admin.AdminOperationException: Error while deleting topic test4
at kafka.admin.TopicCommand$$anonfun$deleteTopic$1.apply(TopicCommand.scala:177)
at kafka.admin.TopicCommand$$anonfun$deleteTopic$1.apply(TopicCommand.scala:162)
at scala.collection.mutable.ResizableArray$class.foreach(ResizableArray.scala:59)
at scala.collection.mutable.ArrayBuffer.foreach(ArrayBuffer.scala:47)
at kafka.admin.TopicCommand$.deleteTopic(TopicCommand.scala:162)
at kafka.admin.TopicCommand$.main(TopicCommand.scala:68)
at kafka.admin.TopicCommand.main(TopicCommand.scala)
(kafka.admin.TopicCommand$)
```

### Possible Causes

The user does not belong to the **kafkaadmin** group. Kafka provides a secure access interface. Only users in the **kafkaadmin** group can delete topics.

### Cause Analysis

1. After the client command is run, the "AdminOperationException" exception is reported.
2. Run the client command **klist** to query the current authenticated user.

```
[root@10-10-144-2 client]# klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: test@HADOOP.COM

Valid starting Expires Service principal
01/25/17 11:06:48 01/26/17 11:06:45 krbtgt/HADOOP.COM@HADOOP.COM
```

The **test** user is used in this example.

3. Run the **id** command to query the user group information.

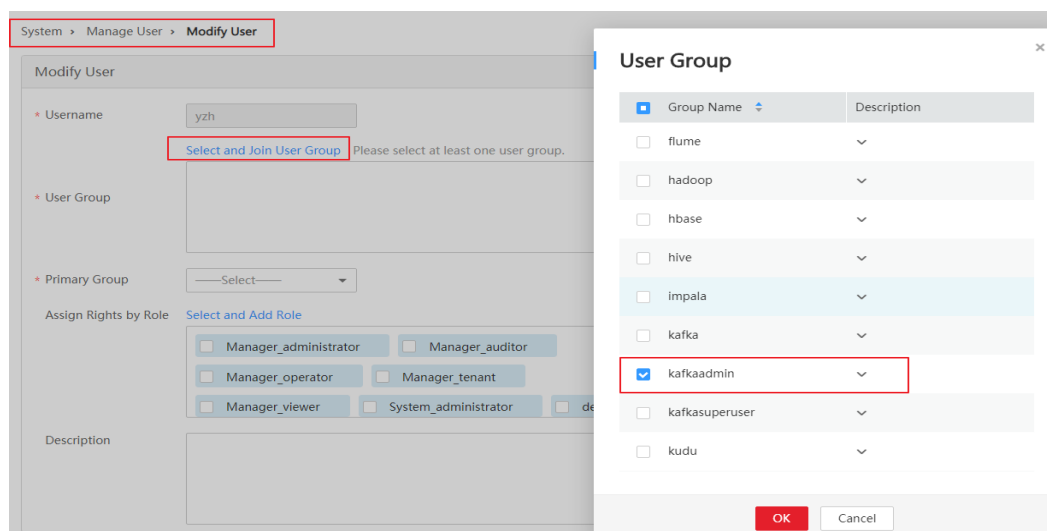
```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop) groups=10001(hadoop),9998(ficommon),10003(kafka)
```

## Solution

MRS Manager:

- Step 1** Log in to MRS Manager.
- Step 2** Choose **System > Manage User**.
- Step 3** In the **Operation** column of the user, click **Modify**.
- Step 4** Add the user to the **kafkaadmin** group. Click **OK**.

**Figure 16-14** Modifying user information



- Step 5** Run the **id** command to query the user group information.

```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop)
groups=10001(hadoop),9998(ficommon),10002(kafkaadmin),10003(kafka)
```

----End

FusionInsight Manager:

- Step 1** Log in to FusionInsight Manager.
- Step 2** Choose **System > Permission > User**.
- Step 3** Locate the row that contains the target user, and click **Modify**.
- Step 4** Add the user to the **kafkaadmin** group. Click **OK**.
- Step 5** Run the **id** command to query the user group information.

```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop)
groups=10001(hadoop),9998(ficommon),10002(kafkaadmin),10003(kafka)
```

----End



## 16.18 When a Kafka Topic Fails to Be Created, "NoAuthException" Is Displayed

### Symptom

When running the following command on the Kafka client to create topics, it is found that the topics cannot be created.

```
kafka-topics.sh --create --zookeeper 192.168.234.231:2181/kafka --replication-factor 1 --partitions 2 --topic test
```

Error messages "NoAuthException" and "KeeperErrorCode = NoAuth for /config/topics" are displayed.

Details are as follows:

```
Error while executing topic command org.apache.zookeeper.KeeperException$NoAuthException:
KeeperErrorCode = NoAuth for /config/topics
org.I0ltec.zkclient.exception.ZkException: org.apache.zookeeper.KeeperException$NoAuthException:
KeeperErrorCode = NoAuth for /config/topics
at org.I0ltec.zkclient.exception.ZkException.create(ZkException.java:68)
at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:685)
at org.I0ltec.zkclient.ZkClient.create(ZkClient.java:304)
at org.I0ltec.zkclient.ZkClient.createPersistent(ZkClient.java:213)
at kafka.utils.ZkUtils$.createParentPath(ZkUtils.scala:215)
at kafka.utils.ZkUtils$.updatePersistentPath(ZkUtils.scala:338)
at kafka.admin.AdminUtils$.writeTopicConfig(AdminUtils.scala:247)
```

### Possible Causes

The user does not belong to the **kafkaadmin** group. Kafka provides a secure access interface. Only users in the **kafkaadmin** group can delete topics.

### Cause Analysis

1. After the client command is run, the "NoAuthException" exception is reported.

```
Error while executing topic command org.apache.zookeeper.KeeperException$NoAuthException:
KeeperErrorCode = NoAuth for /config/topics
org.I0ltec.zkclient.exception.ZkException: org.apache.zookeeper.KeeperException$NoAuthException:
KeeperErrorCode = NoAuth for /config/topics
at org.I0ltec.zkclient.exception.ZkException.create(ZkException.java:68)
at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:685)
at org.I0ltec.zkclient.ZkClient.create(ZkClient.java:304)
at org.I0ltec.zkclient.ZkClient.createPersistent(ZkClient.java:213)
at kafka.utils.ZkUtils$.createParentPath(ZkUtils.scala:215)
at kafka.utils.ZkUtils$.updatePersistentPath(ZkUtils.scala:338)
at kafka.admin.AdminUtils$.writeTopicConfig(AdminUtils.scala:247)
```

2. Run the client command **klist** to query the current authenticated user.

```
[root@10-10-144-2 client]# klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: test@HADOOP.COM

Valid starting Expires Service principal
01/25/17 11:06:48 01/26/17 11:06:45 krbtgt/HADOOP.COM@HADOOP.COM
```

The **test** user is used in this example.

3. Run the **id** command to query the user group information.

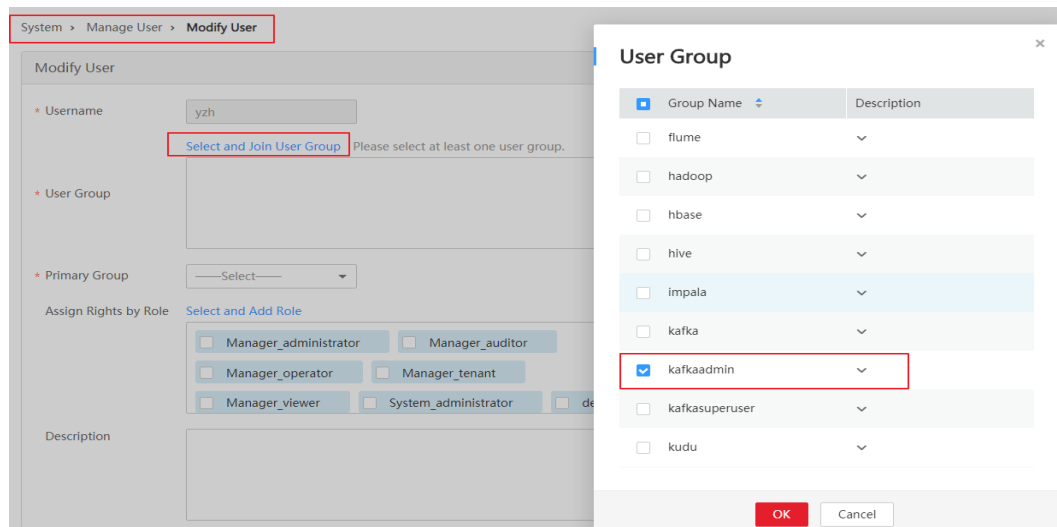
```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop) groups=10001(hadoop),9998(ficommon),10003(kafka)
```

## Solution

MRS Manager:

- Step 1** Log in to MRS Manager.
- Step 2** Choose **System > Manage User**.
- Step 3** In the **Operation** column of the user, click **Modify**.
- Step 4** Add the user to the **kafkaadmin** group.

**Figure 16-15** Adding the user to the **kafkaadmin** group



- Step 5** Run the **id** command to query the user group information.

```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop)
groups=10001(hadoop),9998(ficommon),10002(kafkaadmin),10003(kafka)
```

----End

FusionInsight Manager:

- Step 1** Log in to FusionInsight Manager.
- Step 2** Choose **System > Permission > User**.
- Step 3** Locate the row that contains the target user, and click **Modify**.
- Step 4** Add the user to the **kafkaadmin** group. Click **OK**.
- Step 5** Run the **id** command to query the user group information.

```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop)
groups=10001(hadoop),9998(ficommon),10002(kafkaadmin),10003(kafka)
```

----End

## 16.19 Failed to Set an ACL for a Kafka Topic, and "NoAuthException" Is Displayed

### Symptom

When running the following command on the Kafka client to set the ACL for a topic, it is found that the topic ACL cannot be set.

```
kafka-acls.sh --authorizer-properties zookeeper.connect=10.5.144.2:2181/kafka --topic topic_acl --producer --add --allow-principal User:test_acl
```

The error message "NoAuthException: KeeperErrorCode = NoAuth for /kafka-acl-changes/acl\_changes\_0000000002" is displayed.

Details are as follows:

```
Error while executing ACL command: org.apache.zookeeper.KeeperException$NoAuthException:
KeeperErrorCode = NoAuth for /kafka-acl-changes/acl_changes_0000000002
org.I0ltec.zkclient.exception.ZkException: org.apache.zookeeper.KeeperException$NoAuthException:
KeeperErrorCode = NoAuth for /kafka-acl-changes/acl_changes_0000000002
at org.I0ltec.zkclient.exception.ZkException.create(ZkException.java:68)
at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:995)
at org.I0ltec.zkclient.ZkClient.delete(ZkClient.java:1038)
at kafka.utils.ZkUtils.deletePath(ZkUtils.scala:499)
at kafka.common.ZkNodeChangeNotificationListener$$anonfun$purgeObsoleteNotifications
$1.apply(ZkNodeChangeNotificationListener.scala:118)
at kafka.common.ZkNodeChangeNotificationListener$$anonfun$purgeObsoleteNotifications
$1.apply(ZkNodeChangeNotificationListener.scala:112)
at scala.collection.mutable.ResizableArray$class.foreach(ResizableArray.scala:59)
at scala.collection.mutable.ArrayBuffer.foreach(ArrayBuffer.scala:47)
at
kafka.common.ZkNodeChangeNotificationListener.purgeObsoleteNotifications(ZkNodeChangeNotificati
oner.scala:112)
at kafka.common.ZkNodeChangeNotificationListener.kafka$common$ZkNodeChangeNotificationListener$
$processNotifications(ZkNodeChangeNotificationListener.scala:97)
at
kafka.common.ZkNodeChangeNotificationListener.processAllNotifications(ZkNodeChangeNotificationListe
ner.scala:77)
at kafka.common.ZkNodeChangeNotificationListener.init(ZkNodeChangeNotificationListener.scala:65)
at kafka.security.auth.SimpleAclAuthorizer.configure(SimpleAclAuthorizer.scala:136)
at kafka.admin.AclCommand$.withAuthorizer(AclCommand.scala:73)
at kafka.admin.AclCommand$.addAcl(AclCommand.scala:80)
at kafka.admin.AclCommand$.main(AclCommand.scala:48)
at kafka.admin.AclCommand.main(AclCommand.scala)
Caused by: org.apache.zookeeper.KeeperException$NoAuthException: KeeperErrorCode = NoAuth for /kafka-
acl-changes/acl_changes_0000000002
at org.apache.zookeeper.KeeperException.create(KeeperException.java:117)
at org.apache.zookeeper.KeeperException.create(KeeperException.java:51)
at org.apache.zookeeper.ZooKeeper.delete(ZooKeeper.java:1416)
at org.I0ltec.zkclient.ZkConnection.delete(ZkConnection.java:104)
at org.I0ltec.zkclient.ZkClient$11.call(ZkClient.java:1042)
at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:985)
```

### Possible Causes

The user does not belong to the **kafkaadmin** group. Kafka provides a secure access interface. Only users in the **kafkaadmin** group can perform the setting operation.

## Cause Analysis

1. After the client command is run, the "NoAuthException" exception is reported.
2. Run the client command **klist** to query the current authenticated user.

```
[root@10-10-144-2 client]# klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: test@HADOOP.COM

Valid starting Expires Service principal
01/25/17 11:06:48 01/26/17 11:06:45 krbtgt/HADOOP.COM@HADOOP.COM
```

The **test** user is used in this example.

3. Run the **id** command to query the user group information.

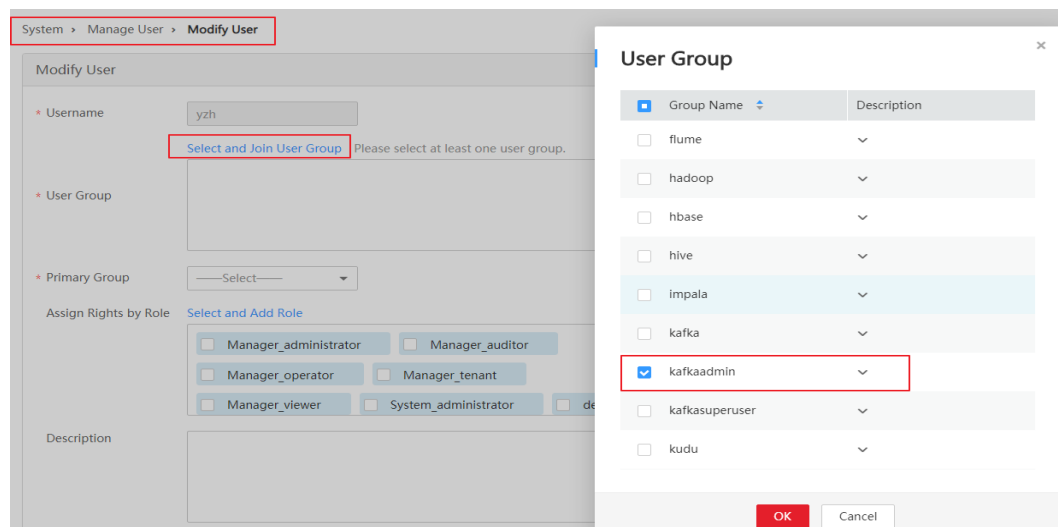
```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop) groups=10001(hadoop),9998(ficommon),10003(kafka)
```

## Solution

MRS Manager:

- Step 1** Log in to MRS Manager.
- Step 2** Choose **System > Manage User**.
- Step 3** In the **Operation** column of the user, click **Modify**.
- Step 4** Add the user to the **kafkaadmin** group.

**Figure 16-16** Modifying a user group



- Step 5** Run the **id** command to query the user group information.

```
[root@host1 client]# id test
uid=20032(test) gid=10001(hadoop)
groups=10001(hadoop),9998(ficommon),10002(kafkaadmin),10003(kafka)
```

----End

FusionInsight Manager:

- Step 1** Log in to FusionInsight Manager.
- Step 2** Choose **System > Permission > User**.

**Step 3** Locate the row that contains the target user, and click **Modify**.

**Step 4** Add the user to the **kafkaadmin** group. Click **OK**.

**Step 5** Run the **id** command to query the user group information.

```
[root@10-10-144-2 client]# id test
uid=20032(test) gid=10001(hadoop)
groups=10001(hadoop),9998(ficcommon),10002(kafkaadmin),10003(kafka)
```

----End

## 16.20 When a Kafka Topic Fails to Be Created, "NoNode for /brokers/ids" Is Displayed

### Symptom

When running the following command on the Kafka client to create topics, it is found that the topics cannot be created.

```
kafka-topics.sh --create --replication-factor 1 --partitions 2 --topic test --zookeeper
192.168.234.231:2181
```

The error message "NoNodeException: KeeperErrorCode = NoNode for /brokers/ids" is displayed.

Details are as follows:

```
Error while executing topic command : org.apache.zookeeper.KeeperException$NoNodeException:
KeeperErrorCode = NoNode for /brokers/ids
[2017-09-17 16:35:28,520] ERROR org.I0ltec.zkclient.exception.ZkNoNodeException:
org.apache.zookeeper.KeeperException$NoNodeException: KeeperErrorCode = NoNode for /brokers/ids
    at org.I0ltec.zkclient.exception.ZkException.create(ZkException.java:47)
    at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:995)
    at org.I0ltec.zkclient.ZkClient.getChildren(ZkClient.java:675)
    at org.I0ltec.zkclient.ZkClient.getChildren(ZkClient.java:671)
    at kafka.utils.ZkUtils.getChildren(ZkUtils.scala:541)
    at kafka.utils.ZkUtils.getSortedBrokerList(ZkUtils.scala:176)
    at kafka.admin.AdminUtils$.createTopic(AdminUtils.scala:235)
    at kafka.admin.TopicCommand$.createTopic(TopicCommand.scala:105)
    at kafka.admin.TopicCommand$.main(TopicCommand.scala:60)
    at kafka.admin.TopicCommand.main(TopicCommand.scala)
Caused by: org.apache.zookeeper.KeeperException$NoNodeException: KeeperErrorCode = NoNode for /
brokers/ids
    at org.apache.zookeeper.KeeperException.create(KeeperException.java:115)
    at org.apache.zookeeper.KeeperException.create(KeeperException.java:51)
    at org.apache.zookeeper.ZooKeeper.getChildren(ZooKeeper.java:2256)
    at org.apache.zookeeper.ZooKeeper.getChildren(ZooKeeper.java:2284)
    at org.I0ltec.zkclient.ZkConnection.getChildren(ZkConnection.java:114)
    at org.I0ltec.zkclient.ZkClient$4.call(ZkClient.java:678)
    at org.I0ltec.zkclient.ZkClient$4.call(ZkClient.java:675)
    at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:985)
    ... 8 more
(kafka.admin.TopicCommand$)
```

### Possible Causes

- The Kafka service is not running.
- The ZooKeeper address parameter in the client command is incorrectly configured.

## Cause Analysis

1. After the client command is run, the "NoNodeException" exception is reported.  
Error while executing topic command : org.apache.zookeeper KeeperException\$NoNodeException: KeeperErrorCode = NoNode for /brokers/ids  
[2017-09-17 16:35:28,520] ERROR org.I0ltec.zkclient.exception.ZkNoNodeException:  
org.apache.zookeeper.KeeperException\$NoNodeException: KeeperErrorCode = NoNode for /brokers/ids  
at org.I0ltec.zkclient.exception.ZkException.create(ZkException.java:47)  
at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:995)  
at org.I0ltec.zkclient.ZkClient.getChildren(ZkClient.java:675)  
at org.I0ltec.zkclient.ZkClient.getChildren(ZkClient.java:671)  
at kafka.utils.ZkUtils.getChildren(ZkUtils.scala:541)  
at kafka.utils.ZkUtils.getSortedBrokerList(ZkUtils.scala:176)  
at kafka.admin.AdminUtils\$.createTopic(AdminUtils.scala:235)  
at kafka.admin.TopicCommand\$.createTopic(TopicCommand.scala:105)  
at kafka.admin.TopicCommand\$.main(TopicCommand.scala:60)  
at kafka.admin.TopicCommand.main(TopicCommand.scala)
2. Check whether the Kafka service is in the normal state on Manager.
3. Check whether the ZooKeeper address in the client command is correct. Check the Kafka information stored in ZooKeeper. The path (Znode) should be suffixed with **/kafka**. It is found that **/kafka** is missing in the configuration.  
[root@10-10-144-2 client]#  
kafka-topics.sh --create --replication-factor 1 --partitions 2 --topic test --zookeeper 192.168.234.231:2181

## Solution

**Step 1** Ensure that the Kafka service is normal.

**Step 2** Add **/kafka** to the ZooKeeper address in the command.

```
[root@10-10-144-2 client]#  
kafka-topics.sh --create --replication-factor 1 --partitions 2 --topic test --zookeeper  
192.168.234.231:2181/kafka
```

----End

## 16.21 When a Kafka Topic Fails to Be Created, "replication factor larger than available brokers" Is Displayed

### Symptom

When running the following command on the Kafka client to create topics, it is found that the topics cannot be created.

```
kafka-topics.sh --create --replication-factor 2 --partitions 2 --topic test --zookeeper  
192.168.234.231:2181
```

The error message "replication factor larger than available brokers" is displayed.

See the following:

```
Error while executing topic command : replication factor: 2 larger than available brokers: 0  
[2017-09-17 16:44:12,396] ERROR kafka.admin.AdminOperationException: replication factor: 2 larger than  
available brokers: 0  
at kafka.admin.AdminUtils$.assignReplicasToBrokers(AdminUtils.scala:117)  
at kafka.admin.AdminUtils$.createTopic(AdminUtils.scala:403)
```

```
at kafka.admin.TopicCommand$.createTopic(TopicCommand.scala:110)
at kafka.admin.TopicCommand$.main(TopicCommand.scala:61)
at kafka.admin.TopicCommand.main(TopicCommand.scala)
(kafka.admin.TopicCommand$)
```

## Possible Causes

- The Kafka service is not running.
- The available Broker of the Kafka service is smaller than the configured **replication-factor**.
- The ZooKeeper address parameter in the client command is incorrectly configured.

## Cause Analysis

1. After the client command is run, "replication factor larger than available brokers" is reported.

```
Error while executing topic command : replication factor: 2 larger than available brokers: 0
[2017-09-17 16:44:12,396] ERROR kafka.admin.AdminOperationException: replication factor: 2 larger
than available brokers: 0
  at kafka.admin.AdminUtils$.assignReplicasToBrokers(AdminUtils.scala:117)
  at kafka.admin.AdminUtils$.createTopic(AdminUtils.scala:403)
  at kafka.admin.TopicCommand$.createTopic(TopicCommand.scala:110)
  at kafka.admin.TopicCommand$.main(TopicCommand.scala:61)
  at kafka.admin.TopicCommand.main(TopicCommand.scala)
(kafka.admin.TopicCommand$)
```

2. Check whether the Kafka service is in the normal state on Manager and whether the current available Broker is smaller than the configured **replication-factor**.
3. Check whether the ZooKeeper address in the client command is correct. Check the Kafka information stored in ZooKeeper. The path (Znode) should be suffixed with **/kafka**. It is found that **/kafka** is missing in the configuration.

```
[root@10-10-144-2 client]#
kafka-topics.sh --create --replication-factor 2 --partitions 2 --topic test --zookeeper
192.168.234.231:2181
```

## Solution

**Step 1** Ensure that the Kafka service is in the normal state and the available Broker is not less than the configured **replication-factor**.

**Step 2** Add **/kafka** to the ZooKeeper address in the command.

```
[root@10-10-144-2 client]#
kafka-topics.sh --create --replication-factor 1 --partitions 2 --topic test --zookeeper
192.168.234.231:2181/kafka
```

----End

## 16.22 Consumer Repeatedly Consumes Data

### Symptom

When the data volume is large, rebalance occurs frequently, causing repeated consumption. The key logs are as follows:

```
2018-05-12 10:58:42,561 | INFO | [kafka-request-handler-3] | [GroupCoordinator 2]: Preparing to restabilize
group DemoConsumer with old generation 118 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
```

```
2018-05-12 10:58:43,245 | INFO | [kafka-request-handler-5] | [GroupCoordinator 2]: Stabilized group DemoConsumer generation 119 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 10:58:43,560 | INFO | [kafka-request-handler-7] | [GroupCoordinator 2]: Assignment received from leader for group DemoConsumer for generation 119 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 10:59:13,562 | INFO | [executor-Heartbeat] | [GroupCoordinator 2]: Preparing to restabilize group DemoConsumer with old generation 119 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 10:59:13,790 | INFO | [kafka-request-handler-3] | [GroupCoordinator 2]: Stabilized group DemoConsumer generation 120 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 10:59:13,791 | INFO | [kafka-request-handler-0] | [GroupCoordinator 2]: Assignment received from leader for group DemoConsumer for generation 120 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 10:59:43,802 | INFO | [kafka-request-handler-2] | Rolled new log segment for '__consumer_offsets-17' in 2 ms. | kafka.log.Log (Logging.scala:68)
2018-05-12 10:59:52,456 | INFO | [group-metadata-manager-0] | [Group Metadata Manager on Broker 2]: Removed 0 expired offsets in 0 milliseconds. | kafka.coordinator.GroupMetadataManager (Logging.scala:68)
2018-05-12 11:00:49,772 | INFO | [kafka-scheduler-6] | Deleting segment 0 from log __consumer_offsets-17. | kafka.log.Log (Logging.scala:68)
2018-05-12 11:00:49,773 | INFO | [kafka-scheduler-6] | Deleting index /srv/BigData/kafka/data4/kafka-logs/__consumer_offsets-17/00000000000000000000.index.deleted | kafka.log.OffsetIndex (Logging.scala:68)
2018-05-12 11:00:49,773 | INFO | [kafka-scheduler-2] | Deleting segment 2147948547 from log __consumer_offsets-17. | kafka.log.Log (Logging.scala:68)
2018-05-12 11:00:49,773 | INFO | [kafka-scheduler-4] | Deleting segment 4282404355 from log __consumer_offsets-17. | kafka.log.Log (Logging.scala:68)
2018-05-12 11:00:49,775 | INFO | [kafka-scheduler-2] | Deleting index /srv/BigData/kafka/data4/kafka-logs/__consumer_offsets-17/0000000002147948547.index.deleted | kafka.log.OffsetIndex (Logging.scala:68)
2018-05-12 11:00:49,775 | INFO | [kafka-scheduler-4] | Deleting index /srv/BigData/kafka/data4/kafka-logs/__consumer_offsets-17/0000000004282404355.index.deleted | kafka.log.OffsetIndex (Logging.scala:68)
2018-05-12 11:00:50,533 | INFO | [kafka-scheduler-6] | Deleting segment 4283544095 from log __consumer_offsets-17. | kafka.log.Log (Logging.scala:68)
2018-05-12 11:00:50,569 | INFO | [kafka-scheduler-6] | Deleting index /srv/BigData/kafka/data4/kafka-logs/__consumer_offsets-17/0000000004283544095.index.deleted | kafka.log.OffsetIndex (Logging.scala:68)
2018-05-12 11:02:21,178 | INFO | [kafka-request-handler-2] | [GroupCoordinator 2]: Preparing to restabilize group DemoConsumer with old generation 120 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 11:02:22,839 | INFO | [kafka-request-handler-4] | [GroupCoordinator 2]: Stabilized group DemoConsumer generation 121 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 11:02:23,169 | INFO | [kafka-request-handler-1] | [GroupCoordinator 2]: Assignment received from leader for group DemoConsumer for generation 121 | kafka.coordinator.GroupCoordinator (Logging.scala:68)
2018-05-12 11:02:49,913 | INFO | [kafka-request-handler-6] | Rolled new log segment for '__consumer_offsets-17' in 2 ms. | kafka.log.Log (Logging.scala:68)
```

In the logs, "Preparing to restabilize group DemoConsumer with old generation" indicates that rebalance occurs.

## Possible Causes

The parameter settings are improper.

## Cause Analysis

**Cause:** Due to improper parameter settings, the data processing time is too long when the data volume is large. Balance frequently occurs, and the offset cannot be submitted normally. As a result, the data is repeatedly consumed.

**Principle:** The offset is submitted only after the poll data is processed. If the processing duration after the poll data is processed exceeds the duration specified by **session.timeout.ms**, the rebalance occurs. As a result, the consumption fails and the offset of the consumed data cannot be submitted. Therefore, the data is consumed at the old offset next time. As a result, the data is repeatedly consumed.



## Solution

Adjust the following service parameters on Manager:

- `request.timeout.ms=100000`
- `session.timeout.ms=90000`
- `max.poll.records=50`
- `heartbeat.interval.ms=3000`

Among the preceding parameters:

- The value of `request.timeout.ms` is 10s greater than that of `session.timeout.ms`.
- The value of `session.timeout.ms` must be within the values of `group.min.session.timeout.ms` and `group.max.session.timeout.ms` on the server.

Set the parameters as required. The `max.poll.records` parameter specifies the number of records for each poll. The purpose is to ensure that the processing time of poll data does not exceed the value of `session.timeout.ms`.

## Related Information

- The post-poll data processing must be efficient and do not block the next poll.
- The poll method and data processing suggestion are processed asynchronously.

# 16.23 Leader for the Created Kafka Topic Partition Is Displayed as none

## Symptom

When a user creates a topic using the Kafka client command, the leader for the created topic partition is displayed as **none**.

```
[root@10-10-144-2 client]#  
kafka-topics.sh --create --replication-factor 1 --partitions 2 --topic test --zookeeper 10.6.92.36:2181/  
kafka
```

```
Created topic "test".
```

```
[root@10-10-144-2 client]#  
kafka-topics.sh --describe --zookeeper 10.6.92.36:2181/kafka  
  
Topic:test    PartitionCount:2    ReplicationFactor:2    Configs:  
Topic: test   Partition: 0    Leader: none    Replicas: 2,3    Isr:  
Topic: test   Partition: 1    Leader: none    Replicas: 3,1    Isr:
```

## Possible Causes

- The Kafka service is not running.
- The user group information cannot be found.

## Cause Analysis

1. Check the Kafka service status:
  - MRS Manager: Log in to MRS Manager and choose **Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Kafka**. Check the Kafka status. The status is **Good**, and the monitoring metrics are correctly displayed.
2. Obtain the Controller node information on the Kafka overview page.
3. Log in to the node where the Controller resides, and run the `cd /var/log/Bigdata/kafka/broker` command to go to the node log directory. The **state-change.log** contains "NoAuthException", which indicates that the ZooKeeper permission is incorrect.

```
2018-05-31 09:20:42,436 | ERROR | [ZkClient-EventThread-34-10.6.92.36:24002,10.6.92.37:24002,10.6.92.38:24002/kafka] | Controller 4 epoch 6 initiated state change for partition [test,1] from NewPartition to OnlinePartition failed | state.change.logger (Logging.scala:103)
```

```
org.I0ltec.zkclient.exception.ZkException: org.apache.zookeeper.KeeperException$NoAuthException: KeeperErrorCode = NoAuth for /brokers/topics/test/partitions
at org.I0ltec.zkclient.exception.ZkException.create(ZkException.java:68)
at org.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:1000)
at org.I0ltec.zkclient.ZkClient.create(ZkClient.java:527)
at org.I0ltec.zkclient.ZkClient.createPersistent(ZkClient.java:293)
```

4. Check on ZooKeeper audit logs recorded in the specified period also indicates that the permission is abnormal.

```
2018-05-31 09:20:42,421 | ERROR | CommitProcWorkThread-1 | session=0xc3000007015d5a18 user=10.6.92.39,kafka/hadoop.hadoop.com@HADOOP.COM,kafka/hadoop.hadoop.com@HADOOP.COM ip=10.6.92.39 operation=create znode target=ZooKeeperServer znode=/kafka/brokers/topics/test/partitions/0/state result=failure
2018-05-31 09:20:42,423 | ERROR | CommitProcWorkThread-1 | session=0xc3000007015d5a18 user=10.6.92.39,kafka/hadoop.hadoop.com@HADOOP.COM,kafka/hadoop.hadoop.com@HADOOP.COM ip=10.6.92.39 operation=create znode target=ZooKeeperServer znode=/kafka/brokers/topics/test/partitions/0 result=failure
2018-05-31 09:20:42,435 | ERROR | CommitProcWorkThread-1 | session=0xc3000007015d5a18 user=10.6.92.39,kafka/hadoop.hadoop.com@HADOOP.COM,kafka/hadoop.hadoop.com@HADOOP.COM ip=10.6.92.39 operation=create znode target=ZooKeeperServer znode=/kafka/brokers/topics/test/partitions result=failure
2018-05-31 09:20:42,439 | ERROR | CommitProcWorkThread-1 | session=0xc3000007015d5a18 user=10.6.92.39,kafka/hadoop.hadoop.com@HADOOP.COM,kafka/hadoop.hadoop.com@HADOOP.COM ip=10.6.92.39 operation=create znode target=ZooKeeperServer znode=/kafka/brokers/topics/test/partitions/1/state result=failure
2018-05-31 09:20:42,441 | ERROR | CommitProcWorkThread-1 | session=0xc3000007015d5a18 user=10.6.92.39,kafka/hadoop.hadoop.com@HADOOP.COM,kafka/hadoop.hadoop.com@HADOOP.COM ip=10.6.92.39 operation=create znode target=ZooKeeperServer znode=/kafka/brokers/topics/test/partitions/1 result=failure
2018-05-31 09:20:42,453 | ERROR | CommitProcWorkThread-1 | session=0xc3000007015d5a18 user=10.6.92.39,kafka/hadoop.hadoop.com@HADOOP.COM,kafka/hadoop.hadoop.com@HADOOP.COM ip=10.6.92.39 operation=create znode target=ZooKeeperServer znode=/kafka/brokers/topics/test/partitions result=failure
```

5. Run the `id -Gn kafka` command on each ZooKeeper instance node. It is found that user group information cannot be queried on a node.

```
[root @bdpsit3ap03 ~]# id -Gn kafka
id: kafka: No such user
[root @bdpsit3ap03 ~]#
```

6. In an MRS cluster, user management is provided by the LDAP service and depends on the SSSD or NSCD service of the OS. The process from creating a user to synchronizing the user to the SSSD service takes some time. If the user does not take effect or the SSSD version has bugs, the user may be invalid on the ZooKeeper node in some cases, which causes topic creation failures.

## Solution

**Step 1** Restart the SSD/NSCD service.

- Red Hat/Euler  
**service sssd restart**
- SUSE  
**sevice nscd restart**

**Step 2** After restarting related services, run the **id Username** command on the node to check whether the user information is valid.

----End

## 16.24 Safety Instructions on Using Kafka

### Brief Introduction to API for Kafka

- New Producer API  
Indicates the API defined in `org.apache.kafka.clients.producer.KafkaProducer`. When `kafka-console-producer.sh` is used, the API is used by default.
- Old Producer API  
Indicates the API defined in `kafka.producer.Producer`. When `kafka-console-producer.sh` is used, the API is invoked to add `--old-producer`.
- New Consumer API  
Indicates the API defined in `org.apache.kafka.clients.consumer.KafkaConsumer`. When `kafka-console-consumer.sh` is used, the API is invoked to add `--new-consumer`.
- Old Consumer API  
Indicates the API defined in `kafka.consumer.ConsumerConnector`. When **`kafka-console-consumer.sh`** is used, the API is used by default.

#### NOTE

New Producer API and new Consumer API are called new API in general in the document.

### Protocol Description for Accessing Kafka

The protocols used to access Kafka are as follows: PLAINTEXT, SSL, SASL\_PLAINTEXT, and SASL\_SSL.

When Kafka is started, access monitoring using the PLAINTEXT and SASL\_PLAINTEXT protocols is started. You can set **`ssl.mode.enable`** to **`true`** in Kafka service configurations to start monitoring using SSL and SASL\_SSL protocols.

The following table describes the four protocols:

Protocol Type	Description	Supported API	Default Port
PLAINTEXT	Supports plaintext access without authentication.	New and old APIs	9092
SASL_PLAINTEXT	Supports plaintext access with Kerberos authentication.	New API	21007
SSL	Supports SSL-encrypted access without authentication.	New API	9093
SASL_SSL	Supports SSL-encrypted access with Kerberos authentication.	New API	21009

## ACL Settings for Topic

Kafka supports secure access. Therefore, users can set the ACL for topics to control that different users access different topics. To view and set the permission information about a topic, run the `kafka-acls.sh` script on the Linux client.

- Scenarios

Assign Kafka users with specific permissions for related topics based on service requirements.

The following table describes default Kafka user groups.

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	Users added to this group have permissions to read data from and write data to all topics.
kafka	Kafka common user group. If users in this group want to read data from and write data to a specific topic, the users in the kafkaadmin group must grant permissions to users in this group.

- Prerequisites

- a. The system administrator has understood service requirements and prepared a Kafka administrator (belonging to the kafkaadmin group).
- b. The Kafka client has been installed.

- Procedure

- a. Log in to the node where the Kafka client is installed as the client installation user.
- b. Switch to the Kafka client installation directory, for example, **/opt/kafkaclient**.  
**cd /opt/kafkaclient**
- c. Run the following command to configure environment variables:  
**source bigdata\_env**
- d. Run the following command to perform user authentication (skip this step for a cluster in common mode):  
**kinit Component service user**
- e. Run the following command to switch to the Kafka client installation directory:  
**cd Kafka/kafka/bin**
- f. The following describes the commands commonly used for user authorization when **kafka-acl.sh** is used:
  - View the permission control list of a topic:  
**./kafka-acls.sh --authorizer-properties  
zookeeper.connect=<ZooKeeper cluster service IP:2181/kafka > --  
list --topic <Topic name>**
  - Add the Producer permission for a user:  
**./kafka-acls.sh --authorizer-properties  
zookeeper.connect=<ZooKeeper cluster service IP:2181/kafka > --  
add --allow-principal User:<username> --producer --topic <Topic  
name>**
  - Remove the Producer permission from a user:  
**./kafka-acls.sh --authorizer-properties  
zookeeper.connect=<ZooKeeper cluster service IP:2181/kafka > --  
remove --allow-principal User:<username> --producer --topic  
<Topic name>**
  - Add the Consumer permission for a user:  
**./kafka-acls.sh --authorizer-properties  
zookeeper.connect=<ZooKeeper cluster service IP:2181/kafka > --  
add --allow-principal User:<username> --consumer --topic <Topic  
name> --group <consumer group name>**
  - Remove the Consumer permission from a user:  
**./kafka-acls.sh --authorizer-properties  
zookeeper.connect=<ZooKeeper cluster service IP:2181/kafka > --  
remove --allow-principal User:<username> --consumer --topic  
<Topic name> --group <consumer group name>**

## Use of New and Old Kafka APIs in Different Scenarios

- Scenario 1: accessing the topic with an ACL

Used API	User Group	Client Parameter	Server Parameter	Access Port
New API	Users need to meet one of the following conditions: <ul style="list-style-type: none"> <li>• In the administrator group</li> <li>• In the <b>kafkaadmin</b> group</li> <li>• In the <b>kafkasuperuser</b> group</li> <li>• In the <b>kafka</b> group and be authorized</li> </ul>	security.protocol=SASL_PLAINTEXT sasl.kerberos.service.name = kafka	-	sasl.port (The default number is 21007.)
		security.protocol=SASL_SSL sasl.kerberos.service.name = kafka	Set ssl.mode.enable to true.	sasl-ssl.port (The default port number is 21009.)
Old API	N/A	N/A	N/A	N/A

- Scenario 2: accessing the topic without an ACL

Used API	User Group	Client Parameter	Server Parameter	Access Port
New API	Users need to meet one of the following conditions: <ul style="list-style-type: none"> <li>• In the administrator group</li> <li>• In the <b>kafkaadmin</b> group</li> <li>• In the <b>kafkasuperuser</b> group</li> </ul>	security.protocol=SASL_PLAINTEXT sasl.kerberos.service.name = kafka	-	sasl.port (The default number is 21007.)
	Users are in the <b>kafka</b> group.		Set <b>allow.everyone.if.no.acl.found</b> to <b>true</b> .	sasl.port (The default number is 21007.)
	Users need to meet one of the following conditions: <ul style="list-style-type: none"> <li>• In the administrator group</li> <li>• In the <b>kafkaadmin</b> group</li> <li>• In the <b>kafkasuperuser</b> group</li> </ul>	security.protocol=SASL_SSL sasl.kerberos.service.name = kafka	Set <b>ssl-enable</b> to <b>true</b> .	sasl-ssl.port (The default port number is 21009.)
	Users are in the <b>kafka</b> group.		Set <b>allow.everyone.if.no.acl.found</b> to <b>true</b> . Set <b>ssl-enable</b> to <b>true</b> .	sasl-ssl.port (The default port number is 21009.)

Used API	User Group	Client Parameter	Server Parameter	Access Port
	-	security.protocol=PLAINTEXT	Set <b>allow.everyone.if.no.acl.found to true.</b>	port (The default number is 21005.)
	-	security.protocol=SSL	Set <b>allow.everyone.if.no.acl.found to true.</b> Set <b>ssl-enable to true.</b>	ssl.port (The default number is 21008.)
Old Producer	-	-	Set <b>allow.everyone.if.no.acl.found to true.</b>	port (The default number is 21005.)
Old Consumer	-	-	Set <b>allow.everyone.if.no.acl.found to true.</b>	ZooKeeper service port: clientPort (The default number is 24002.)

## 16.25 Obtaining Kafka Consumer Offset Information

### Symptom

How do I obtain Kafka Consumer offset information when using Kafka Consumer to consume data?

### Kafka APIs

- New Producer API  
Indicates the API defined in **org.apache.kafka.clients.producer.KafkaProducer**. When **kafka-console-producer.sh** is used, the API is used by default.
- Old Producer API  
Indicates the API defined in **kafka.producer.Producer**. When **kafka-console-producer.sh** is used, the API is invoked to add **--old-producer**.
- New Consumer API  
Indicates the API defined in **org.apache.kafka.clients.consumer.KafkaConsumer**. When **kafka-console-consumer.sh** is used, the API is invoked to add **--new-consumer**.
- Old Consumer API



Indicates the API defined in **kafka.consumer.ConsumerConnector**. When **kafka-console-consumer.sh** is used, the API is used by default.

#### NOTE

New Producer API and new Consumer API are called new API in general in the document.

## Procedure

### Old Consumer API

- Prerequisites
  - a. The system administrator has understood service requirements and prepared a Kafka administrator (belonging to the kafkaadmin group).
  - b. The Kafka client has been installed.
- Procedure
  - a. Log in to the node where the Kafka client is installed as the client installation user.
  - b. Switch to the Kafka client installation directory, for example, **/opt/kafkaclient**.  
**cd /opt/kafkaclient**
  - c. Run the following command to configure environment variables:  
**source bigdata\_env**
  - d. Run the following command to perform user authentication (skip this step for a cluster in common mode):  
**kinit Component service user**
  - e. Run the following command to switch to the Kafka client installation directory:  
**cd Kafka/kafka/bin**
  - f. Run the following command to obtain Consumer offset metric information:

```
bin/kafka-consumer-groups.sh --zookeeper <zookeeper_host:port>/kafka --list
```

```
bin/kafka-consumer-groups.sh --zookeeper <zookeeper_host:port>/kafka --describe --group test-consumer-group
```

Example:

```
kafka-consumer-groups.sh --zookeeper 192.168.100.100:2181/kafka --list  
kafka-consumer-groups.sh --zookeeper 192.168.100.100:2181/kafka --describe --group test-consumer-group
```

### New Consumer API

- Prerequisites
  - a. The system administrator has understood service requirements and prepared a Kafka administrator (belonging to the kafkaadmin group).
  - b. The Kafka client has been installed.
- Procedure
  - a. Log in to the node where the Kafka client is installed as the client installation user.

- b. Switch to the Kafka client installation directory, for example, `/opt/client`.  
**cd /opt/client**
- c. Run the following command to configure environment variables:  
**source bigdata\_env**
- d. Run the following command to perform user authentication (skip this step for a cluster in common mode):  
**kinit Component service user**
- e. Run the following command to switch to the Kafka client installation directory:  
**cd Kafka/kafka/bin**
- f. Run the following command to obtain Consumer offset metric information:  
**kafka-consumer-groups.sh --bootstrap-server <broker\_host:port> --describe --group my-group**  
Example:  
**kafka-consumer-groups.sh --bootstrap-server 192.168.100.100:9092 --describe --group my-group**

## 16.26 Adding or Deleting Configurations for a Topic

### Symptom

Configure or modify a specific topic when using Kafka.

Parameters that can be modified at the topic level:

```
cleanup.policy  
compression.type  
delete.retention.ms  
file.delete.delay.ms  
flush.messages  
flush.ms  
index.interval.bytes  
max.message.bytes  
min.cleanable.dirty.ratio  
min.insync.replicas  
preallocate  
retention.bytes  
retention.ms  
segment.bytes  
segment.index.bytes  
segment.jitter.ms  
segment.ms  
unclean.leader.election.enable
```

### Procedure

- Prerequisites  
The Kafka client has been installed.
- Procedure
  - a. Log in to the node where the Kafka client is installed as the client installation user.

- b. Switch to the Kafka client installation directory, for example, `/opt/client`.  
**cd /opt/client**
- c. Run the following command to configure environment variables:  
**source bigdata\_env**
- d. Run the following command to perform user authentication (skip this step for a cluster in common mode):  
**kinit Component service user**
- e. Run the following command to switch to the Kafka client installation directory:  
**cd Kafka/kafka/bin**
- f. Run the following commands to configure and delete a topic:  
**kafka-topics.sh --alter --topic <topic\_name> --zookeeper <zookeeper\_host:port>/kafka --config <name=value>**  
**kafka-topics.sh --alter --topic <topic\_name> --zookeeper <zookeeper\_host:port>/kafka --delete-config <name>**  
Example:  
**kafka-topics.sh --alter --topic test1 --zookeeper 192.168.100.100:2181/kafka --config retention.ms=86400000**  
**kafka-topics.sh --alter --topic test1 --zookeeper 192.168.100.100:2181/kafka --delete-config retention.ms**
- g. Run the following command to query topic information:  
**kafka-topics.sh --describe -topic <topic\_name> --zookeeper <zookeeper\_host:port>/kafka**

## 16.27 Reading the Content of the `__consumer_offsets` Internal Topic

### Issue

The user does not know how does Kafka save the offset of a consumer to the `__consumer_offsets` of internal topics.

### Procedure

- Step 1** Log in to the node where the Kafka client is installed as the client installation user.
- Step 2** Switch to the Kafka client installation directory, for example, `/opt/client`.  
**cd /opt/client**
- Step 3** Run the following command to configure environment variables:  
**source bigdata\_env**
- Step 4** Run the following command to perform user authentication (skip this step for a cluster in common mode):  
**kinit Component service user**

**Step 5** Run the following command to switch to the Kafka client installation directory:

```
cd Kafka/kafka/bin
```

**Step 6** Run the following command to obtain Consumer offset metric information:

```
kafka-console-consumer.sh --topic __consumer_offsets --zookeeper  
<zk_host:port>/kafka --formatter  
"kafka.coordinator.group.GroupMetadataManager\  
$OffsetsMessageFormatter" --consumer.config <property file> --from-  
beginning
```

Add the following content to the *<property file>* configuration file:

```
exclude.internal.topics = false
```

Example:

```
kafka-console-consumer.sh --topic __consumer_offsets --zookeeper  
10.5.144.2:2181/kafka --formatter  
"kafka.coordinator.group.GroupMetadataManager\  
$OffsetsMessageFormatter" --consumer.config ../config/consumer.properties  
--from-beginning
```

```
[example-group1, test2, 0]::[OffsetMetadata[0, NO_METADATA], CommitTime 1487121209218, ExpirationTime 14872076  
9218]  
[example-group1, test2, 1]::[OffsetMetadata[0, NO_METADATA], CommitTime 1487121209218, ExpirationTime 14872076  
9218]  
[example-group1, test2, 0]::[OffsetMetadata[2, NO_METADATA], CommitTime 1487121269208, ExpirationTime 14872076  
9208]  
[example-group1, test2, 1]::[OffsetMetadata[1, NO_METADATA], CommitTime 1487121269208, ExpirationTime 14872076  
9208]
```

----End

## 16.28 Configuring Logs for Shell Commands on the Kafka Client

### Issue

How do I set the log level for shell commands on the client?

### Procedure

**Step 1** Log in to the node where the Kafka client is installed as the client installation user.

**Step 2** Switch to the Kafka client installation directory, for example, */opt/client*.

```
cd /opt/client
```

**Step 3** Run the following command to switch to the Kafka client configuration directory:

```
cd Kafka/kafka/config
```

**Step 4** Open the *tools-log4j.properties* file, change **WARN** to **INFO**, and save the file.

```
log4j.rootLogger=WARN, stderr  
  
log4j.appender.stderr=org.apache.log4j.ConsoleAppender  
log4j.appender.stderr.layout=org.apache.log4j.PatternLayout  
log4j.appender.stderr.layout.ConversionPattern=[%d] %p %m (%c)%n  
log4j.appender.stderr.Target=System.err
```

```
log4j.rootLogger=INFO, stderr
log4j.appender.stderr=org.apache.log4j.ConsoleAppender
log4j.appender.stderr.layout=org.apache.log4j.PatternLayout
log4j.appender.stderr.layout.ConversionPattern=[%d] %p %m (%c)%n
log4j.appender.stderr.Target=System.err
```

**Step 5** Switch to the Kafka client installation directory, for example, **/opt/client**.

```
cd /opt/client
```

**Step 6** Run the following command to configure environment variables:

```
source bigdata_env
```

**Step 7** Run the following command to perform user authentication (skip this step for a cluster in common mode):

```
kinit Component service user
```

**Step 8** Run the following command to switch to the Kafka client installation directory:

```
cd Kafka/kafka/bin
```

**Step 9** Run the following command to obtain the topic information. The log information can be viewed on the console.

```
kafka-topics.sh --list --zookeeper 10.5.144.2:2181/kafka
[2017-02-17 14:34:27,005] INFO JAAS File name: /opt/client/Kafka/./kafka/config/jaas.conf
(org.I0ltec.zkclient.ZkClient)
[2017-02-17 14:34:27,007] INFO Starting ZkClient event thread. (org.I0ltec.zkclient.ZkEventThread)
...
Test
__consumer_offsets
counter
test
test2
test3
test4
```

```
----End
```

## 16.29 Obtaining Topic Distribution Information

### Issue

How do I obtain topic distribution information in a Broker instance?

### Preparations

- Prerequisites  
The Kafka and ZooKeeper clients have been installed.
- Procedure
  - a. Log in to the node where the Kafka client is installed as the client installation user.
  - b. Switch to the Kafka client installation directory, for example, **/opt/client**.  

```
cd /opt/client
```

- c. Run the following command to configure environment variables:  
**source bigdata\_env**
- d. Run the following command to perform user authentication (skip this step for a cluster in common mode):  
**kinit Component service user**
- e. Run the following command to switch to the Kafka client installation directory:  
**cd Kafka/kafka/bin**
- f. Run the Kafka commands to obtain the topic assignment information and copy synchronization information, and check the return result.  
**kafka-topics.sh --describe --zookeeper <zk\_host:port/chroot>**

Example:

```
[root@mgt-dat-sh-3-01-3 client]#kafka-topics.sh --describe --zookeeper 10.149.0.90:2181/kafka
Topic:topic1 PartitionCount:2 ReplicationFactor:2 Configs:
Topic: topic1 Partition: 0 Leader: 26 Replicas: 23,25 Isr: 26
Topic: topic1 Partition: 1 Leader: 24 Replicas: 24,23 Isr: 24,23
```

In the preceding information, **Replicas** indicates the replica assignment information and **Isr** indicates the replica synchronization information.

## Solution 1

1. Query the Broker ID mapping in ZooKeeper.  
**sh zkCli.sh -server <zk\_host:port>**
2. Run the following command on the ZooKeeper client:  
**ls /kafka/brokers/ids**  
**get/kafka/brokers/ids/<queried Broker ID>**

Example:

```
[root@node-master1gAMQ kafka]# zkCli.sh -server node-master1gAMQ:2181
Connecting to node-master1gAMQ:2181
Welcome to ZooKeeper!
JLine support is enabled

WATCHER::

WatchedEvent state:SyncConnected type:None path:null
[zk: node-master1gAMQ:2181(CONNECTED) 0] ls /kafka/brokers/ids seqid topics
[zk: node-master1gAMQ:2181(CONNECTED) 0] ls /kafka/brokers/ids [1]
[zk: node-master1gAMQ:2181(CONNECTED) 1] get /kafka/brokers/ids/1
{"listener_security_protocol_map":{"PLAINTEXT":"PLAINTEXT","SSL":"SSL"},"endpoints":["PLAINTEXT://192.168.2.242:9092","SSL://192.168.2.242:9093"],"rack":"/default/rack0","jmx_port":21006,"host":"192.168.2.242","timestamp":"1580886124398","port":9092,"version":4}
[zk: node-master1gAMQ:2181(CONNECTED) 2]
```

## Solution 2

Obtain the mapping between nodes and Broker IDs.

**kafka-broker-info.sh --zookeeper <zk\_host:port/chroot>**

Example:

```
[root@node-master1gAMQ kafka]# bin/kafka-broker-info.sh --zookeeper 192.168.2.70:2181/kafka
Broker_ID IP_Address
```

---

1 192.168.2.242

## 16.30 Kafka HA Usage Description

### Kafka High Reliability and Availability

Kafka message transmission assurance mechanism ensures message transmission after required parameters are set to meet different performance and reliability requirements.

- **Kafka high availability and high performance**

If HA and high performance are required, configure parameters listed in the following table.

Parameter	Default Value	Description
unclean.leader.election.enable	true	Specifies whether a replica that is not in the ISR can be selected as the leader. If this parameter is set to <b>true</b> , data may be lost.
auto.leader.rebalance.enable	true	Specifies whether the leader automated balancing function is used.  If this parameter is set to <b>true</b> , the controller periodically balances the leader of each partition on all nodes and assigns the leader to a replica with a higher priority.

Parameter	Default Value	Description
acks	1	<p>The leader needs to check whether the message has been received and determine whether the required operation has been processed. This parameter affects message reliability and performance.</p> <ul style="list-style-type: none"> <li>• If this parameter is set to <b>0</b>, the Producer does not wait for any response from the server and the message is considered successful.</li> <li>• If this parameter is set to <b>1</b>, when the leader of the copy verifies that data has been written into the cluster, the leader makes repose quickly without waiting until all the copies are written. In this case, if the leader is abnormal when the leader makes the confirmation but replica synchronization is not complete, data will be lost.</li> <li>• If this parameter is set to <b>-1</b> (all), the synchronization is successful only after all synchronization copies are confirmed. If <b>min.insync.replicas</b> is also configured, multiple copies can be written successfully. In this case, as long as one copy remains active, the record is not lost.</li> </ul> <p><b>NOTE</b> This parameter is configured in the Kafka client configuration file.</p>
min.insync.replicas	1	Specifies the minimum number of replicas to which data is written when <b>acks</b> is set to <b>-1</b> for the Producer.

Impact of HA and high performance configurations:

**NOTICE**

After HA and high performance are configured, the data reliability decreases. Specifically, data may be lost of disks or nodes are faulty.

- **Kafka high reliability configuration**

If high data reliability is required, configure parameters listed in the following table.



Parameter	Recommended Value	Description
unclean.leader.election.enable	false	Indicates whether a replica that is not in the ISR list can be elected as a leader.
acks	-1	<p>The leader needs to check whether the message has been received and determine whether the required operation has been processed.</p> <p>If this parameter is set to <b>-1</b>, the message is successfully received only when all replicas in the ISR list have confirmed to receive the message. The <b>min.insync.replicas</b> parameter must also be set to ensure that multiple copies can be written successfully. As long as one copy is active, the record is not lost.</p> <p><b>NOTE</b> This parameter is configured in the Kafka client configuration file.</p>
min.insync.replicas	2	<p>Specifies the minimum number of replicas to which data is written when <b>acks</b> is set to <b>-1</b> for the Producer.</p> <p>Ensure that the value of <b>Min.insync.replicas</b> is equal to or less than that of <b>replication.factor</b>.</p>

Impact of high reliability configurations:

- Deteriorated performance  
All copies in the ISR list are required, and the writing of the minimum number of copies has been verified successful. As a result, the delay of a single message increases and the processing capability of the client decreases. The actual performance depends on the onsite test data.
- Reduced availability  
A replica that is not in the ISR list cannot be elected as a leader. If the leader goes offline and other replicas are not in the ISR list, the partition remains unavailable until the leader node recovers.  
All copies in the ISR list are required, and the writing of the minimum number of copies has been verified successful. When the node where a copy of a partition is located is faulty, the minimum number of successful copies cannot be met. As a result, service writing fails.

## Configuration Impact

Evaluate reliability and performance requirements based on service scenarios and use proper parameter configuration.

 NOTE

- For valuable data, you are advised to configure raid1 or raid5 for Kafka data directory disks to improve data reliability in case disk fault of a single disk.
- The **acks** parameter is named different for different Producer APIs.
  - New Producer API  
Indicates the interface defined in **org.apache.kafka.clients.producer.KafkaProducer**. The **acks** parameter name remains unchanged for this API.
  - Old Producer API  
Indicates the interface defined in **kafka.producer.Producer**. The **acks** parameter is named as **request.required.acks** for this API.
- For parameters that can be modified at the topic level, the service level configurations are used by default. These parameters can be separately configured based on topic reliability requirements.  
For example, you can configure the reliability parameters of the topic named **test**.  
**kafka-topics.sh --zookeeper 192.168.1.205:2181/kafka --alter --topic test --config unclean.leader.election.enable=false --config min.insync.replicas=2**  
**192.168.1.205** indicates the ZooKeeper service IP address.
- If modification of the service-level requires the restart of Kafka, you are advised to modify the service-level configuration on the change page.

## 16.31 Failed to Manage a Kafka Cluster Using the Kafka Shell Command

### Symptom

A user fails to run the Kafka command to view the current topic information.

The error information is as follows:

```
root@Slave2bin]#./kafka-topics.sh --describe --topic example-metric1 --zookeeper
192.119.147.231:2181,192.119.147.228:2181,192.119.147.227:2181
Error while executing topic command:org.apache.zookeeper.KeeperException
$AuthFailedException:KeeperErrorCode=AuthFailedfor/brokers/topics
[2016-09-26 16:58:59,873]ERRORorg.I0ltec.zkclient.exception.ZkException:org.apache.zookeeper.KeeperExcept
ion$AuthFailedException:KeeperErrorCode=AuthFailedfor/brokers/topics
atorg.I0ltec.zkclient.exception.ZkException.create(ZkException.java:68)
atorg.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:995)
atorg.I0ltec.zkclient.ZkClient.getChildren(ZkClient.java:675)
atorg.I0ltec.zkclient.ZkClient.getChildren(ZkClient.java:671)
atkafka.utils.ZkUtils.getChildrenParentMayNotExist(ZkUtils.scala:548)
atkafka.utils.ZkUtils.getAllTopics(ZkUtils.scala:798)
atkafka.admin.TopicCommand$.getTopics(TopicCommand.scala:82)
atkafka.admin.TopicCommand$.describeTopic(TopicCommand.scala:183)
atkafka.admin.TopicCommand$.main(TopicCommand.scala:66)
atkafka.admin.TopicCommand.main(TopicCommand.scala)
Causedby:org.apache.zookeeper.KeeperException$AuthFailedException:KeeperErrorCode=AuthFailedfor/
brokers/topics
atorg.apache.zookeeper.KeeperException.create(KeeperException.java:127)
atorg.apache.zookeeper.KeeperException.create(KeeperException.java:51)
atorg.apache.zookeeper.ZooKeeper.getChildren(ZooKeeper.java:2256)
atorg.apache.zookeeper.ZooKeeper.getChildren(ZooKeeper.java:2284)
atorg.I0ltec.zkclient.ZkConnection.getChildren(ZkConnection.java:114)
atorg.I0ltec.zkclient.ZkClient$.call(ZkClient.java:678)
atorg.I0ltec.zkclient.ZkClient$.call(ZkClient.java:675)
atorg.I0ltec.zkclient.ZkClient.retryUntilConnected(ZkClient.java:985)
```

```
...8more  
(kafka.admin.TopicCommand$)
```

## Cause Analysis

The user has checked the permission of the account for running this command. The account has the highest permission to perform operations for Kafka. This problem is not caused by lack of permission.

The command executed is incorrect. It should contain **/kafka** to access the path (Znode) of the Kafka information stored on ZooKeeper. The correct command is as follows:

```
root@Slave2bin]#./kafka-topics.sh --describe --topic example-metric1 --zookeeper  
192.168.147.231:2181,192.168.147.228:2181,192.168.147.227:2181/kafka
```

## Solution

Use the correct command to query topic information. For details, see [Shell](#).

# 16.32 Kafka Producer Writes Oversized Records

## Symptom

When a user develops a Kafka application and invokes the new interface (**org.apache.kafka.clients.producer.\***) as a Producer to write data to Kafka, the size of a single record is 1100055, which exceeds the value (**1000012**) of **message.max.bytes** in the Kafka configuration file **server.properties**. After the values of **message.max.bytes** and **replica.fetch.max.bytes** in the Kafka service configurations are changed to **5242880**, the exception persists.

The following error information is displayed:

```
.....  
14749 [Thread-0] INFO com.xxx.bigdata.kafka.example.NewProducer - The ExecutionException occurred : {}.  
java.util.concurrent.ExecutionException: org.apache.kafka.common.errors.RecordTooLargeException: The  
message is 1100093 bytes when serialized which is larger than the maximum request size you have  
configured with the max.request.size configuration.  
at org.apache.kafka.clients.producer.KafkaProducer$FutureFailure.<init>(KafkaProducer.java:739)  
at org.apache.kafka.clients.producer.KafkaProducer.doSend(KafkaProducer.java:483)  
at org.apache.kafka.clients.producer.KafkaProducer.send(KafkaProducer.java:430)  
at org.apache.kafka.clients.producer.KafkaProducer.send(KafkaProducer.java:353)  
at com.xxx.bigdata.kafka.example.NewProducer.run(NewProducer.java:150)  
Caused by: org.apache.kafka.common.errors.RecordTooLargeException: The message is **** bytes when  
serialized which is larger than the maximum request size you have configured with the max.request.size  
configuration.  
.....
```

## Cause Analysis

When data is written to Kafka, the Kafka client compares the value of **max.request.size** with the size of the data to be written. If the size of the data to be written exceeds the default value of **max.request.size**, the preceding exception is reported.

## Solution

**Step 1** You can set the value of **max.request.size** when initializing the Kafka Producer instance.

For example, you can set this parameter to **5252880** as follows:

```
// Protocol type: Currently, the SASL_PLAINTEXT or PLAINTEXT protocol types can be used.
props.put(securityProtocol, kafkaProc.getValues(securityProtocol, "SASL_PLAINTEXT"));
// Service name
props.put(saslKerberosServiceName, "kafka");
props.put("max.request.size", "5252880");
.....
```

----End

## 16.33 Kafka Consumer Reads Oversized Records

### Symptom

After data is written to Kafka, a user develops an application and invokes the interface (**org.apache.kafka.clients.consumer.\***) to read data from Kafka as a Consumer. However, the reading fails and the following error is reported:

```
.....
1687 [KafkaConsumerExample] INFO org.apache.kafka.clients.consumer.internals.AbstractCoordinator -
Successfully joined group DemoConsumer with generation 1
1688 [KafkaConsumerExample] INFO org.apache.kafka.clients.consumer.internals.ConsumerCoordinator -
Setting newly assigned partitions [default-0, default-1, default-2] for group DemoConsumer
2053 [KafkaConsumerExample] ERROR com.xxx.bigdata.kafka.example.NewConsumer -
[KafkaConsumerExample], Error due to
org.apache.kafka.common.errors.RecordTooLargeException: There are some messages at [Partition=Offset]:
{default-0=177} whose size is larger than the fetch size 1048576 and hence cannot be ever returned.
Increase the fetch size on the client (using max.partition.fetch.bytes), or decrease the maximum message
size the broker will allow (using message.max.bytes).
2059 [KafkaConsumerExample] INFO com.xxx.bigdata.kafka.example.NewConsumer -
[KafkaConsumerExample], Stopped
.....
```

### Cause Analysis

When reading data, the Kafka client compares the size of the data to be read with the value of **max.partition.fetch.bytes**. If the size exceeds the value of **max.partition.fetch.bytes**, the preceding exception is reported.

### Solution

**Step 1** When creating a Kafka Consumer instance during initialization, set **max.partition.fetch.bytes**.

For example, you can set this parameter to **5252880** as follows:

```
.....
// Security protocol type
props.put(securityProtocol, kafkaProc.getValues(securityProtocol, "SASL_PLAINTEXT"));
// Service name
props.put(saslKerberosServiceName, "kafka");

props.put("max.partition.fetch.bytes", "5252880");
.....
```

----End

## 16.34 High Usage of Multiple Disks on a Kafka Cluster Node

### Issue

The usage of multiple disks on a node in the Kafka streaming cluster is high. The Kafka service will become unavailable if the usage reaches 100%.

### Symptom

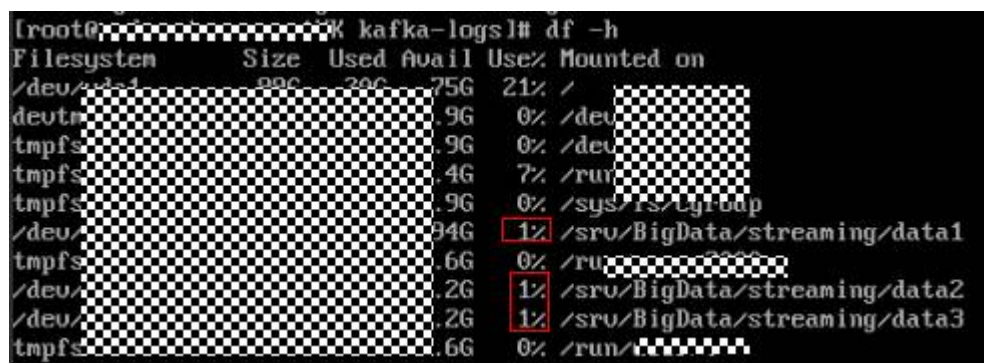
A node in the MRS Kafka streaming cluster purchased by the customer has multiple disks. Due to improper partitioning and service reasons, the usage of some disks is high. When the usage reaches 100%, Kafka becomes unavailable.

### Cause Analysis

The disk data needs to be processed in a timely manner. After the value of **log.retention.hours** is changed, the service needs to be restarted. To ensure service continuity, you can shorten the aging time of a single data-intensive topic as required.

### Procedure

- Step 1** Log in to the core node of the Kafka streaming cluster.
- Step 2** Run the **df -h** command to check the disk usage.



```
[root@kafka-logs1 ~]# df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       80G   20G   60G  21% /
devtmpfs        9G     0B   9G   0% /dev
tmpfs           9G     0B   9G   0% /dev
tmpfs           4G     0B   4G   0% /run
tmpfs           9G     0B   9G   0% /sys/fs/cgroup
/dev/sda2       84G   12G   72G   1% /sru/BigData/streaming/data1
tmpfs           6G     0B   6G   0% /run
/dev/sda3       2G     0B   2G   0% /sru/BigData/streaming/data2
/dev/sda4       2G     0B   2G   0% /sru/BigData/streaming/data3
tmpfs           6G     0B   6G   0% /run
```

- Step 3** Obtain the data storage directory from the **log.dirs** configuration item in the Kafka configuration file **/opt/Bigdata/MRS\_2.1.0/1\_11\_Broker/etc/server.properties**.

Change the configuration file path based on the cluster version in the environment. If there are multiple disks, use commas (,) to separate multiple configuration items.

```
ssl.port = 9093
log.dirs = /srv/BigData/streaming/data1/kafka-logs,/srv/BigData/streaming/data2/kafka-logs,/srv/BigData/streaming/data3/kafka-logs
controlled.shutdown.enable = true
compression.type = producer
max.connections.per.ip.overrides =
log.message.timestamp.difference.max.ms = 9223372036854775807
sasl.kerberos.kinit.cmd = /opt/BigData/MRS_2.1.0/install/FusionInsight-kerberos-1.15.2/kerberos/bin/kinit
log.cleaner.io.max.bytes.per.second = 1.7976931348623157E308
auto.leader.rebalance.enable = true
leader.inbalance.check.interval.seconds = 300
log.cleaner.min.cleanable.ratio = 0.5
```

- Step 4** Run the `cd` command to go to the data storage directory obtained in **Step 3** of the disk with high usage.
- Step 5** Run the `du -sh *` command to print the name and size of the current topic.

```
[root@node-str-coreethK kafka-logs]# du -sh *
0      offset-checkpoint
12K    0
4.0K   0-t-offset-checkpoint
4.0K   0-properties
4.0K   0-point-offset-checkpoint
4.0K   0-0-offset-checkpoint
20K    0-t-0
20K    0-t-1
20K    0-t-2
20K    0-t-3
20K    0-t-4
20K    0-t-5
[root@node-str-coreethK kafka-logs]# pwd
/srv/BigData/streaming/data1/kafka-logs
```

```
[root@node-str-coreethK kafka-logs]# du -sh *
0      0-r-offset-checkpoint
4.0K   0-art-offset-checkpoint
4.0K   0-properties
4.0K   0-ry-point-offset-checkpoint
4.0K   0-ation-offset-checkpoint
4.0K   0-0
4.0K   0-1
4.0K   0-2
4.0K   0-6
4.0K   0-8
[root@node-str-coreethK kafka-logs]# pwd
/srv/BigData/streaming/data2/kafka-logs
```

```
[root@node-str-coreethK kafka-logs]# du -sh *
0      0-r-offset-checkpoint
4.0K   0-art-offset-checkpoint
4.0K   0-properties
4.0K   0-ry-point-offset-checkpoint
4.0K   0-ation-offset-checkpoint
4.0K   0-3
4.0K   0-4
4.0K   0-5
4.0K   0-7
4.0K   0-9
[root@node-str-coreethK kafka-logs]# pwd
/srv/BigData/streaming/data3/kafka-logs
```

**Step 6** Determine the method of changing the data retention period. The default global data retention period of Kafka is seven days. A large amount of data may be written to some topics, and these topics reside on the partitions on the disk with high usage.

- You can change the global data retention period to a smaller value to release disk space. This method requires a Kafka service restart, which may affect service running. For details, see [Step 7](#).
- You can change the data retention period of a single topic to a smaller value to release disk space. This configuration takes effect without a Kafka service restart. For details, see [Step 8](#).

**Step 7** Log in to Manager. On the Kafka service configuration page, switch to **All Configurations** and search for the **log.retention.hours** configuration item. The default value is 7 days. Change it based on the site requirements.

**Step 8** Change the data retention time of the topics on these disks.

1. Check the retention time of the topic data.

```
bin/kafka-topics.sh --describe --zookeeper <ZooKeeper cluster service IP address>:2181/kafka --topic kctest
```

```
lroot@node-master:~/kafka# bin/kafka-topics.sh --describe --zookeeper 192.168.201.175:2181/kafka --topic kctest
Topic:kctest   PartitionCount:1   ReplicationFactor:1   Configs:retention.ms:1000000
Topic: kctest   Partition: 0       Leader: 1               Replicas: 1           Isr: 1
```

2. Set the topic data retention time. **--topic** indicates the topic name, and **retention.ms** indicates the data retention time, in milliseconds.

```
kafka-topics.sh --zookeeper <ZooKeeper cluster service IP address>:2181/
kafka --alter --topic kctest --config retention.ms=1000000
```

```
lroot@node-master:~/kafka# kafka-topics.sh --zookeeper 192.168.201.175:2181/kafka --alter --topic kctest --config retention.ms=1000000
WARNING: Altering topic configuration from this script has been deprecated and may be removed in future releases.
        Going forward, please use kafka-configs.sh for this functionality
Updated config for topic "kctest".
```

After the data retention time is set, the deletion operation may not be performed immediately. The deletion operation starts after the time specified by **log.retention.check.interval.ms**. You can check whether the **delete** field exists in the **server.log** file of Kafka to determine whether the deletion operation takes effect. If the **delete** field exists, the deletion operation has taken effect. You can also run the **df -h** command to check the disk usage and determine whether the setting takes effect.

```
log.retention.check.interval.ms = 300000
```

----End

## 16.35 Kafka Is Disconnected from the ZooKeeper Client

### Symptom

The Kafka client is disconnected from the ZooKeeper client.

Error message "Received event :disconnected::none: with disconnected Writer Zookeeper" is displayed in the client log.

## Cause Analysis

When the Kafka client connects to the ZooKeeper client, the fsync operation takes a long time. As a result, the connection times out.

## Procedure

**Step 1** Log in to the node where the ZooKeeper client is located.

```
cd Client installation directory
```

```
source bigdata_env
```

```
kinit Component service user (Skip this step for clusters with Kerberos authentication disabled.)
```

**Step 2** Run the following command to modify the file:

```
vim Client installation directory/zookeeper/conf/zoo.cfg
```

Increase the values of **tickTime** and **syncLimit** in the file.

For example, set **tickTime** to **3000** and **syncLimit** to **7**.

**Step 3** Save the file.

```
----End
```



# 17 Using Oozie

## 17.1 Oozie Jobs Do Not Run When a Large Number of Jobs Are Submitted Concurrently

### Symptom

When a large number of Oozie jobs are submitted concurrently, the jobs do not run.

### Cause Analysis

When Oozie submits a job, an oozie-launcher job is started first, and then the oozie-launcher job submits the real job for execution. By default, the oozie-launcher job and the real job are in the same queue.

When a large number of Oozie jobs are submitted concurrently, a large number of oozie-launcher jobs may be started, exhausting the resources of the queue. As a result, no more resources are available to start real jobs, and the jobs are not executed.

### Procedure

- Step 1** Create a queue for Oozie by referring to **Creating a Tenant** in the *User Guide*. Alternatively, you can directly use the **launcher-job** queue generated during MRS cluster creation.
- Step 2** On Manager, choose **Cluster > Services > Oozie > Configurations**, search for **oozie.site.configs**, and add **oozie.launcher.default.queue** as the parameter name and the queue name as the value, for example, **launcher-job**.

Parameter	Value	Description	Parameter File
oozie-oozozoo			
oozie.processing.timezone	UTC	[[Desc]] Oozie server timezone. Valid values are UTC and GMT[+ -]####. For ex...	oozie/oozie-site.xml
oozie.mll.connector.port	21002	[[Desc]]_mll.connector.port_ [Default] 21002 [Range] 21002-21004	oozie/oozie-site.xml
oozie.mll.registry.port	21002	[[Desc]]_mll.registry.port_ [Default] 21002 [Range] 21002-21004	oozie/oozie-site.xml
oozie.service.HadoopAccessorService.supported filesystems	*	[[Desc]]List the different filesystems supported for federation. If wildcard "*" is...	hadoop/core-site.xml
oozie.site.configs	oozie.launcher.default.queue launcher.job	[[Desc]]Add a customized configuration item to the global file oozie-site.xml	oozie/oozie-site.xml

----End

## 17.2 An Error Is Reported When Oozie Schedules HiveSQL Jobs

### Symptom

In an MRS 3.x cluster, Oozie fails to schedule Hive jobs. The HiveSQL logs show that the scheduling task is executed successfully but the Yarn job fails. The error information is as follows:

```
java.io.IOException:output.properties data exceeds its limit [2048]
```

### Cause Analysis

A large number of jobs are submitted in the HiveSQL script at a time, and the size of the job data exceeds the maximum value (2 KB, 2048 bytes) allowed by Oozie Launcher. Therefore, you need to increase the default value. Adjusting this parameter value does not affect the cluster performance.

### Procedure

- Step 1** Log in to Manager, choose **Cluster > Services > Oozie > Configurations > All Configurations > Customization**, and add parameter **oozie.action.max.output.data** and value **204800** to the **oozie-site.xml** file.

	Name	Value	
oozie.site.configs	oozie.action.max.output.data	204800	+

- Step 2** Save the settings and restart the Oozie service.

- Step 3** Execute the Oozie scheduling job again.

----End

## 17.3 Oozie Tasks Cannot Be Submitted on a Client Outside the MRS Cluster or Can Be Submitted Only Two Hours Later

### Symptom

A user on a client outside the MRS cluster failed to submit Oozie tasks or successfully submitted them only two hours later.

### Cause Analysis

Some background logs about the Java secure random number are found. In the JDK, the SecureRandom algorithm depends on the random data provided by the OS. In the Linux OS, **/dev/random** and **/dev/urandom** store data related to the secure random number. When the entropy pool is empty, read operations

from **/dev/random** are blocked until the entropy pool collects enough environment noise data. As a result, task submission becomes slow or fails.

## Procedure

**Step 1** Run the following command on the node where the client outside the cluster is located to install rng-tools:

```
yum install rng-tools
```

**Step 2** Run the following command to start rng-tools:

```
systemctl start rngd
```

To enable rng-tools to automatically run after its node is restarted, run the following command:

```
systemctl enable rngd
```

**Step 3** Run the Oozie task again.

**----End**

# 18 Using Presto

## 18.1 During sql-standard-with-group Configuration, a Schema Fails to Be Created and the Error Message "Access Denied" Is Displayed

### Issue

A schema fails to be created during sql-standard-with-group configuration and the error message "Access Denied" is displayed.

### Symptom

The following error is reported:

```
CREATE SCHEMA hive.sf2 WITH (location = 'obs://obs-zy1234/sf2');Query 20200224_031203_00002_g6gzy failed: Access Denied: Cannot create schema sf2
```

### Cause Analysis

Creating a schema in Presto requires the Hive administrator permission. The current user does not have the permission.

### Procedure

MRS Manager:

- Method 1:
  - a. Log in to MRS Manager and choose **System > Manage User**.
  - b. Locate the row that contains the target user, and click **Modify** in the **Operation** column.
  - c. Click **Select and Add Role** to assign the **System\_administrator** permission to the user.
  - d. Click **OK**.
- Method 2:

- a. Log in to MRS Manager and choose **System > Manage Role**.
- b. Click **Create Role** and set the following parameters:
  - Enter a role name, for example, **hive\_admin**.
  - Set **Permission** to **Hive** and select **Hive Admin Privilege**.
- c. Click **OK** to save the role.
- d. Choose **System > Manage User**.
- e. Locate the row that contains the target user, and click **Modify** in the **Operation** column.
- f. Click **Select and Add Role** to add the newly created **hive\_admin** permission to the user.
- g. Click **OK**.

FusionInsight Manager:

- Method 1:
  - a. Log in to FusionInsight Manager and choose **System > Permission > User**.
  - b. Locate the row that contains the target user, and click **Modify** in the **Operation** column.
  - c. Click **Add** next to the role to assign the **System\_administrator** permission to the user.
  - d. Click **OK**.
- Method 2:
  - a. Log in to FusionInsight Manager and choose **System > Permission > Role**.
  - b. Click **Create Role** and set the following parameters:
    - Enter a role name, for example, **hive\_admin**.
    - To configure resource permissions, select **Hive** and **Hive Admin Permissions**.
  - c. Click **OK** to save the role.
  - d. Choose **System > Permission > User**.
  - e. Locate the row that contains the target user, and click **Modify** in the **Operation** column.
  - f. Click **Add** next to the role to add the **hive\_admin** permission for the user.
  - g. Click **OK**.

## 18.2 Coordinator Process of Presto Cannot Be Started

### Issue

The coordinator process of Presto is terminated due to an unknown reason or cannot be started.

## Symptom

The Presto coordinator process cannot be started properly. On the Manager page, it is shown that the presto coordinator process is started properly and its status is normal. However, the background log shows that the coordinator process is not started. Only the following log is displayed:

```

2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.config-spec
null
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.environment
null
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.internal-address-source
IP
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.location
null
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.bind-ip
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.external-address
null
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.id
Coordinator-
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.internal-address
2020-06-18T18:17:02.872+0800 INFO main Bootstrap node.pool
general
2020-06-18T18:20:00.014+0800 INFO main in-airlift-log-logging Disabling stderr output
2020-06-18T18:20:01.777+0800 INFO main Bootstrap PROPERTY
DEFAULT RUNTIME
DESCRIPTION
2020-06-18T18:20:01.777+0800 INFO main Bootstrap event.max-output-stage-size
16MB 16MB
2020-06-18T18:20:01.777+0800 INFO main Bootstrap query.client.timeout
5.00m 5.00m
2020-06-18T18:20:01.777+0800 INFO main Bootstrap query.initial-hash-partitions
100 32
2020-06-18T18:20:01.777+0800 INFO main Bootstrap query.manager.initialization-required-workers
1 1
Minimum number of workers that must be available before the cluster will accept queries
2020-06-18T18:20:01.777+0800 INFO main Bootstrap query.manager.initialization-timeout
5.00m 5.00m
After this time, the cluster will accept queries even if the minimum required workers are not available
2020-06-18T18:20:01.778+0800 INFO main Bootstrap query.max-concurrent-queries
1000 1000
2020-06-18T18:20:01.778+0800 INFO main Bootstrap query.max-history
100 100
@@@

```

The Presto coordinator process is terminated before being started, and no other logs are printed. Other Presto logs do not indicate the reason.

## Cause Analysis

The port check logic of the Presto health check script does not distinguish ports.

## Procedure

**Step 1** Use a tool to log in to the master nodes of the cluster and perform the following operations:

**Step 2** Run the following command to edit the file:

```
vim /opt/Bigdata/MRS_xxx/install/FusionInsight-Presto-*/ha/module/harm/plugin/script/pcd.sh
```

Change line 31 in the file to `http_port_exists=$(netstat -apn | awk '{print $4, $6}' | grep :${HTTP_PORT} | grep LISTEN | wc -l)`.

```

25
26 check_status()
27 {
28     proc_exists=$(ps -ef | grep com.facebook.presto.server.PrestoServer | grep -v grep | wc -l)
29     param="-u $PRESTO_SERVER/v1/cluster"
30     if [[ $(proc_exists) == 1 ]]; then
31         http_port_exists=$(netstat -apn | awk '{print $4, $6}' | grep :${HTTP_PORT} | grep LISTEN | wc -l)
32     fi
33     if [[ $(http_port_exists) == 1 ]]; then
34         log ${PCD_LOG_FILE} "INFO" "return [ normal ]"
35         return 0
36     else
37         log ${PCD_LOG_FILE} "ERROR" "HTTP PORT does not exist, return [ abnormal ]"
38         return 2
39     fi
40 else
41     log ${PCD_LOG_FILE} "INFO" " coordinator process not exists, return [ abnormal ]"
42     return 2
43 fi
44 }
45

```

**Step 3** Save the modification. On Manager, choose **Services > Presto > Instances** to restart the Coordinator process.

----End

## 18.3 When Presto Queries a Kudu Table, an Error Is Reported Indicating That the Table Cannot Be Found

### Issue

An error is reported when Presto is used to query a Kudu table.

### Symptom

When Presto is used to query a Kudu table, the following error message is displayed.

```
presto:default> show tables;
      Table
impala::default.kudu_taobao
impala::default.kudu_tt
impala::default.kudutest
(3 rows)

Query 20210201_030636_00026_95mzd, FINISHED, 4 nodes
Splits: 53 total, 53 Done (100.00%)
0:00 [3 rows, 125B] [18 rows/s, 766B/s]

presto:default> select count(*) from kudu.default.kudu_taobao;
Query 20210201_030653_00027_95mzd failed: line 1:22: Table kudu.default.kudu_taobao does not exist
select count(*) from kudu.default.kudu_taobao

presto:default> select count(*) from kudu.taobao;
Query 20210201_030939_00028_95mzd failed: line 1:22: Table kudu.default.kudu_taobao does not exist
select count(*) from kudu.taobao

presto:default>
```

### Error information

```
2021-02-01T15:08:13.850+0800 INFO query-execution-10 io.prestosql.event.QueryMonitor TIMELINE: Query 20210201_070813_00007_5x
9q9 :: Transaction:[72fad2d9-8480-4435-ac8d-ac2a93bf181d] :: elapsed 71ms :: planning 15ms :: waiting 0ms :: scheduling 56ms :: running
1ms :: finishing 0ms :: begin 2021-02-01T15:08:13.736+08:00 :: end 2021-02-01T15:08:13.801+08:00
2021-02-01T15:14:17.487+0800 INFO query-execution-19 io.prestosql.event.QueryMonitor TIMELINE: Query 20210201_071417_00008_5x
9q9 :: Transaction:[0104571a-3ec6-4013-b7c6-0219916a07ba] :: elapsed 369ms :: planning 167ms :: waiting 3ms :: scheduling 45ms :: running
98ms :: finishing 72ms :: begin 2021-02-01T15:14:17.095+08:00 :: end 2021-02-01T15:14:17.464+08:00
2021-02-01T15:15:11.127+0800 INFO query-execution-20 io.prestosql.event.QueryMonitor TIMELINE: Query 20210201_071510_00009_5x
9q9 :: Transaction:[0dc0e086-5500-4932-a528-693cb4d0054] :: elapsed 282ms :: planning 115ms :: waiting 0ms :: scheduling 30ms :: running
95ms :: finishing 82ms :: begin 2021-02-01T15:15:10.830+08:00 :: end 2021-02-01T15:15:11.112+08:00
2021-02-01T15:15:14.006+0800 ERROR remote-task-callback-40 io.prestosql.execution.StageStateMachine Stage 20210201_071513_00
010_6x9q9_1 failed
java.lang.IllegalArgumentException: No page sink provider for catalog 'kudu'
    at com.google.common.base.Preconditions.checkNotNull(Preconditions.java:216)
    at io.prestosql.split.PageSinkManager.providerFor(PageSinkManager.java:67)
    at io.prestosql.split.PageSinkManager.createPageSink(PageSinkManager.java:61)
    at io.prestosql.operator.TableWriterOperatorsTableWriterOperatorFactory.createPageSink(TableWriterOperator.java:114)
    at io.prestosql.operator.TableWriterOperatorsTableWriterOperatorFactory.createOperator(TableWriterOperator.java:105)
    at io.prestosql.operator.DriverFactory.createDriver(DriverFactory.java:114)
    at io.prestosql.execution.SqlTaskExecution$DriverSplitRunnerFactory.createDriver(SqlTaskExecution.java:941)
    at io.prestosql.execution.executor.PrioritizedSplitRunner.processFor(SqlTaskExecution.java:1059)
    at io.prestosql.execution.executor.TaskExecutorsTaskRunner.run(TaskExecutor.java:484)
    at io.prestosql.$gen.Presto_EI_PrestoSQL_Kernel_Component_0_3_308_0100_B001_13_gbcbbafe_dirty_20210201_070255_1.run(Unknown S
ource)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:748)
```

### Cause Analysis

There are no Kudu configurations on the actually running node (node where the worker instance is located).

## Procedure

**Step 1** Add configuration file **kudu.properties** to all worker instance nodes in the Presto cluster.

Path for storing the configuration file: **/opt/Bigdata/MRS\_xxx/1\_x\_Worker/etc/catalog/** (Change the path based on the actual cluster version.)

Configuration file content:

```
connector.name=kudu
kudu.client.master-addresses=KuduMasterIP1:port,KuduMasterIP2:port,KuduMasterIP3:port
```

### NOTE

- Set the IP address and port number of the KuduMaster node based on the site requirements.
- Add the file permission and owner group that are the same as those of other files in the file save path to the configuration file.

**Step 2** After the modification, choose **Components > Kudu** on the cluster details page, click **More**, and select **Restart Service**.

----End

## 18.4 No Data is Found in the Hive Table Using Presto

### Issue

When Presto is used to query a Hive table, no data is found.

### Symptom

Presto cannot query the data written by **union** statements executed by the Tez engine.

### Cause Analysis

When Hive uses the Tez engine to execute the **union** statements, the output file is stored in the **HIVE\_UNION\_SUBDIR** directory. However, Presto does not access files in child directories by default. Therefore, data in the **HIVE\_UNION\_SUBDIR** directory is not read.

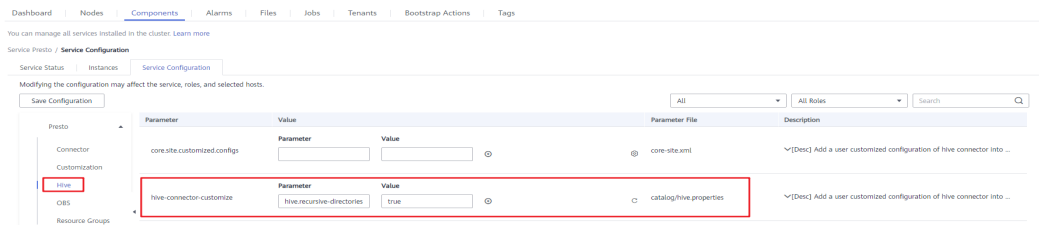
## Procedure

**Step 1** On the MRS console, click the cluster name, and choose **Components > Presto > Service Configuration**.

**Step 2** Change **Basic** to **All**.

**Step 3** In the navigation pane on the left, choose **Presto > Hive**. In the **catalog/hive.properties** file, add the **hive.recursive-directories** parameter and set it to **true**.





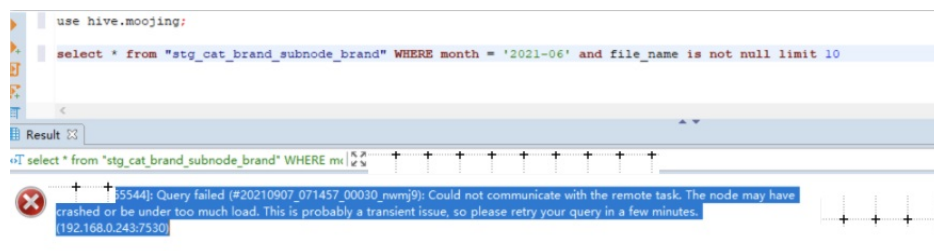
**Step 4** Click **Save Configuration** and select **Restart the affected services or instances**.  
----End

## 18.5 Error Message "The node may have crashed or be under too much load" Is Displayed During MRS Presto Query

### Issue

Error message "The node may have crashed or be under too much load" is displayed when a Presto query statement is executed.

**Figure 18-1** Error information



### Cause Analysis

1. Log in to the master node of the cluster and search the Presto log in the **/var/log/Bigdata** directory.
2. Check the process logs of the Presto worker instance on the corresponding node based on the error information.

The log information is as follows:  
java.lang.OutOfMemoryError: Java heap space

3. According to the error information, it is predicted that out of memory (OOM) occurs. As a result, the query error is reported.

### Procedure

**Step 1** On Manager, search for the JVM parameter (**xmx**) of the worker instance.

**Step 2** Change **1024** next to **xmx** to **2048** and save the configuration.

**Step 3** Restart the Presto service and try again.

----End

## 18.6 Accessing Presto from an MRS Cluster Through a Public Network

### Issue

To debug the Presto JDBC sample code, the user needs to use the public network to access Presto.

### Cause Analysis

Log in to MRS Manager and check all Presto configurations and Coordinator role configurations.

PRESTO\_COORDINATOR\_FLOAT\_IP = *Internal IP address*

Log in to the active node with the Coordinator role and check whether the IP address is provided by **eth0:PRESTO**.

```
[root@node-master2n0Xb ~]# ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.244 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:d5:a0:e8 txqueuelen 4000 (Ethernet)
    RX packets 30206419 bytes 11537075854 (10.7 GiB)
    RX errors 0 dropped 0 overruns 0 frame 7052
    TX packets 26042255 bytes 10021677809 (10.0 GiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

eth0:DBS: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.88 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:d5:a0:e8 txqueuelen 4000 (Ethernet)

eth0:HUE: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.160 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:d5:a0:e8 txqueuelen 4000 (Ethernet)

eth0:JHS: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.38 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:d5:a0:e8 txqueuelen 4000 (Ethernet)

eth0:PRESTO: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.204 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:d5:a0:e8 txqueuelen 4000 (Ethernet)
```

Presto service port: **7520** for a non-security cluster and **7521** for a security cluster

To access Presto through a public network, bind a public network address to the floating NIC and enter the correct JDBC URL.

URL format:

**jdbc:presto://example\_ip:7520/Catalog/schema**

**jdbc:presto://example\_ip:7521/Catalog/schema**

### Procedure

- Step 1** On MRS Manager, locate the Presto component and record its internal floating IP address.
- Step 2** On the VPC console, create an EIP, locate the subnet of the MRS cluster, locate the floating IP address of Presto, and bind an EIP to the floating IP address.

**Step 3** In the security group of the MRS cluster, allow the access from the source address to the Presto port of MRS and test the access. (The following uses a non-security cluster as an example. The catalog is **hive**.)

```
[root@node-master1mnm ~]# presto_cli.sh --server http://119.3.164.55:7520 --catalog hive --schema bigdata
--server http://119.3.164.55:7520 --catalog hive --schema bigdata
presto:bigdata> select * from eimrs;
id | name | worker | happies | age
-----+-----+-----+-----+-----
0  |      |        |         | 30
1  |      |        |         | 28
2  |      |        |         | 32
(3 rows)

Query 20210731_050300_00007_p4y4b, FINISHED, 1 node
Splits: 19 total, 19 done (100.00%)
0:00 [3 rows, 96B] [10 rows/s, 326B/s]

presto:bigdata> █
```

----End

# 19 Using Spark

## 19.1 An Error Is Reported When the Split Size Is Changed for a Running Spark Application

### Issue

An error occurs when the split size is changed in a Spark application.

### Symptom

A user needs to implement multiple mappers by changing the maximum split size to make the Spark application run faster. However, an error occurs when the user runs the **set** command to modify the Hive configuration.

```
0: jdbc:hive2://192.168.1.18:21066/> set mapred.max.split.size=1000000;  
Error: Error while processing statement: Cannot modify mapred.max.split.size at runtime. It is not in list of  
params that are allowed to be modified at runtime( state=42000,code=1)
```

### Cause Analysis

- Before the **hive.security.whitelist.switch** parameter is set to enable or disable the whitelist in security mode, the allowed parameters must have been configured in **hive.security.authorization.sqlstd.confwhitelist**.

The screenshot shows a list of configuration parameters with their values. The parameter `hive.security.whitelist.switch` is highlighted with a blue border and has its value set to `ON`. Other parameters include `hive.security.temporary.function.need.admin` (true), `hive.security.transform.disallow` (true), and `hive.server2.enable.doAs` (true).

- The default whitelist does not contain the **mapred.max.split.size** parameter. Therefore, the system displays a message indicating that the maximum split size cannot be changed.

## Procedure

**Step 1** Go to the Hive configuration page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console, choose **Components** > **Hive** > **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 3.x or later: Log in to FusionInsight Manager and choose **Cluster** > **Services** > **Hive** > **Configurations** > **All Configurations**.

**Step 2** Search for **hive.security.authorization.sqlstd.confwhitelist.append**, and add **mapred.max.split.size** to **hive.security.authorization.sqlstd.confwhitelist.append**. For details, see [Using Hive from Scratch](#).

**Step 3** Save the configuration and restart the Hive component.

**Step 4** Run the **set mapred.max.split.size=1000000**; command on the Hive Beeline CLI again.

----End

## 19.2 Incorrect Parameter Format Is Displayed When a Spark Task Is Submitted

### Symptom

A user fails to submit a cluster task using Spark and a message is displayed indicating that the parameter format is incorrect.

```
[omninode-master1-qxvMQ spark]$
[omninode-master1-qxvMQ spark]$
[omninode-master1-qxvMQ spark]$
[omninode-master1-qxvMQ spark]$ ./bin/spark-submit --class cn.interf.Test --master yarn-client /opt/client/Spark/spark1-1.0-SNAPSHOT.jar:
Error: Unrecognized option: --class cn.interf.Test --master

Java HotSpot(TM) 64-Bit Server VM warning: Cannot open file <LOG_DIR>/gc.log due to No such file or directory

Usage: spark-submit [options] <app jar | python file> [app arguments]
Usage: spark-submit --kill [submission ID] --master [spark://...]
Usage: spark-submit --status [submission ID] --master [spark://...]
Usage: spark-submit run-example [options] example-class [example args]

Options:
--master MASTER_URL      spark://host:port, mesos://host:port, yarn, or local.
--deploy-mode DEPLOY_MODE  whether to launch the driver program locally ("client") or
                           on one of the worker machines inside the cluster ("cluster")
                           (Default: client).
--class CLASS_NAME        Your application's main class (for Java / Scala apps).
--name NAME              A name of your application.
--jars JARS              Comma-separated list of local jars to include on the driver
```

### Cause Analysis

- The command executed contains invalid characters.
- The owner and owner group of the uploaded JAR package are incorrect.

## Procedure

- Step 1** Run `./bin/spark-submit --class cn.interf.Test --master yarn-client Client installation directory/Spark/spark1-1.0-SNAPSHOT.jar`; to check whether invalid characters are imported.
- Step 2** If they are imported, modify the invalid characters and run the command again.
- Step 3** After the command is executed again, other errors occur. Both the owner and the owner group of the JAR package are **root**.
- Step 4** Change the owner and the owner group of the JAR package to **omm:wheel**.

----End

## 19.3 Spark, Hive, and Yarn Are Unavailable Due to Insufficient Disk Capacity

### Issue

A critical alarm indicating that Spark, Hive, and Yarn are unavailable is generated due to insufficient disk capacity.

### Symptom

The disk capacity of the cluster created by the user is insufficient. As a result, a critical alarm indicating that Spark, Hive, and Yarn are unavailable is generated.

### Cause Analysis

Insufficient cluster disk capacity affects HDFS data writing. When the HDFS disk space usage exceeds the threshold, HDFS becomes abnormal. As a result, Spark, Hive, and Yarn are unavailable.

The alarm indicating that Spark, Hive, and Yarn are unavailable is generated due to insufficient disk capacity of the cluster. After the disk capacity is expanded, the alarm is cleared. Therefore, it can be determined that the HDFS function fault is caused by insufficient disk capacity.

### Procedure

For details about how to clear the alarm triggered by insufficient disk capacity, see [ALM-12017 Insufficient Disk Capacity](#).

### Related Information

For details about how to solve the problem that HDFS disk usage exceeds the threshold, see [ALM-14001 HDFS Disk Usage Exceeds the Threshold](#).

## 19.4 A Spark Job Fails to Run Due to Incorrect JAR File Import

### Symptom

A Spark job fails to be executed.

### Cause Analysis

The imported JAR file is incorrect when the Spark job is executed. As a result, the Spark job fails to be executed.

### Procedure

**Step 1** Log in to any Master node.

**Step 2** Run the `cd /opt/Bigdata/MRS_*/install/FusionInsight-Spark-*/spark/examples/jars` command to view the JAR file of the sample program.

#### NOTE

A JAR file name contains a maximum of 1023 characters and cannot include special characters (;|&>,<'\$). In addition, it cannot be left blank or full of spaces.

**Step 3** Check the path of the executable programs in HDFS or the OBS bucket. The path may vary depending on the file system.

- OBS storage path: starts with `obs://`, for example, `obs://wordcount/program/hadoop-mapreduce-examples-2.7.x.jar`.
- HDFS storage path: starts with `/user`. The Spark Script must end with `.sql`, and MapReduce and Spark must end with `.jar`. `.sql` and `.jar` are case-insensitive.

----End

## 19.5 Spark Job Suspended Due to Insufficient Memory or Lack of JAR Packages

### Issue

The memory is insufficient or no JAR package is added when a Spark job is submitted. As a result, the job is in the pending state for a long time or memory overflow occurs during job running.

### Symptom

The job is pending for a long time after being submitted. The following error information is displayed after the job is executed repeatedly:

```
Exception in thread "main" org.apache.spark.SparkException: Job aborted due to stage failure:
Aborting TaskSet 3.0 because task 0 (partition 0) cannot run anywhere due to node and executor blacklist.
Blacklisting behavior can be configured via spark.blacklist.*.
```

## Cause Analysis

The memory is insufficient or no JAR package is added when the job is submitted.

## Procedure

**Step 1** Check whether the JAR package is added when the job is submitted.

- If yes, go to [Step 2](#).
- If no, add the JAR package. If the job execution becomes normal, no further action is required. If the job is still in the pending state for a long time, go to [Step 2](#).

**Step 2** Log in to the MRS console, click a cluster name on the **Active Clusters** page and view the node specifications of the cluster on the **Nodes** page.

**Step 3** Add cluster resources for the **nodemanager** process.

Operations on MRS Manager:

1. Log in to MRS Manager and choose **Services > Yarn > Service Configuration**.
2. Set **Type** to **All**, and then search for **yarn.nodemanager.resource.memory-mb** in the search box to view the value of this parameter. You are advised to set the parameter value to 75% to 90% of the total physical memory of nodes.

Operations on FusionInsight Manager:

1. Log in to FusionInsight Manager. Choose **Cluster > Services > Yarn**.
2. Choose **Configurations > All Configurations**. Search for **yarn.nodemanager.resource.memory-mb** in the search box and check the parameter value. You are advised to set the parameter value to 75% to 90% of the total physical memory of nodes.

**Step 4** Modify the Spark service configuration.

Operations on MRS Manager:

1. Log in to MRS Manager and choose **Services > Spark > Service Configuration**.
2. Set **Type** to **All**, and then search for **spark.driver.memory** and **spark.executor.memory** in the search box.

Set these parameters to a larger or smaller value based on the complexity and memory requirements of the submitted Spark job. (Generally, the values need to be increased.)

Operations on FusionInsight Manager:

1. Log in to FusionInsight Manager. Choose **Cluster > Services > Spark**.
2. Choose **Configurations > All Configurations**. Search for **spark.driver.memory** and **spark.executor.memory** in the search box and increase or decrease the values based on actual requirements. Generally, increase the values based on the complexity and memory of the submitted Spark job.



 NOTE

- If a SparkJDBC job is used, search for **SPARK\_EXECUTOR\_MEMORY** and **SPARK\_DRIVER\_MEMORY** and change their values based on the complexity and memory requirements of the submitted Spark job. (Generally, the values need to be increased.)
- If the number of cores needs to be specified, you can search for **spark.driver.cores** and **spark.executor.cores** and change their values.

**Step 5** Scale out the cluster if the preceding requirements still cannot be met because Spark depends on the memory for computing.

----End

## 19.6 Error "ClassNotFoundException" Is Reported When a Spark Task Is Submitted

### Symptom

When a Spark task is executed, an error message is displayed, indicating that class **ClassNotFoundException** cannot be found.

The error message is as follows:

```
Exception encountered | org.apache.spark.internal.Logging$class.logError(Logging.scala:91)  
org.apache.hadoop.hbase.DoNotRetryIOException: java.lang.ClassNotFoundException:  
org.apache.phoenix.filter.SingleCQKeyValueComparisonFilter
```

### Cause Analysis

The default path configured by the user is incorrect.

### Procedure

**Step 1** Log in to any Master node.

**Step 2** Modify the configuration file in the Spark client directory.

Run the **vim *Client installation directory*/Spark/spark/conf/spark-defaults.conf** command to open the **spark-defaults.conf** file and set **spark.executor.extraClassPath** to **\${PWD}/\***.

**Step 3** Submit the task again.

----End

# 19.7 Driver Displays a Message Indicating That the Running Memory Exceeds the Threshold When a Spark Task Is Submitted

## Symptom

A Spark task fails to be submitted because the running memory exceeds the threshold.

## Cause Analysis

The Driver log prints that the applied Executor memory exceeds the cluster limit.  
... INFO Client: Verifying our application has not requested more than the maximum memory capability of the cluster (6144 MB per container)  
... ERROR SparkContext: Error initializing SparkContext.  
java.lang.IllegalArgumentException: Required executor memory (10240+1024 MB) is above the max threshold (6144 MB) of this cluster!

Spark tasks are submitted to Yarn and the resources used by the Executor to run tasks are managed by Yarn. From the error message, you can see that when a user starts the Executor, 10 GB memory is specified, which exceeds the upper memory limit of each Container set by Yarn. As a result, the task cannot be started.

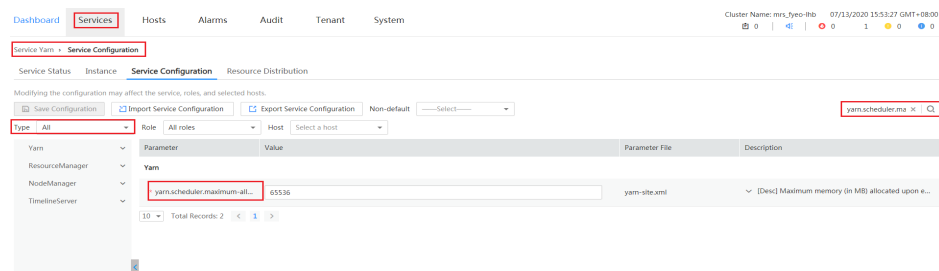
## Solution

Modify the Yarn configuration to increase the restriction on containers. For example, you can adjust parameter **yarn.scheduler.maximum-allocation-mb** to control the resources for starting the Executor. Restart the Yarn service after the modification.

MRS Manager:

- Step 1** Log in to MRS Manager.
- Step 2** Choose **Services > Yarn > Service Configuration** and set **Type** to **All**.
- Step 3** In **Search**, enter **yarn.scheduler.maximum-allocation-mb** to modify the parameter, save the configuration, and then restart the service. See the following figure.

Figure 19-1 Modifying Yarn service parameters



----End

FusionInsight Manager:

- Step 1** Log in to FusionInsight Manager.
  - Step 2** Choose **Cluster > Service > Yarn**. Click **Configurations** and select **All Configurations**.
  - Step 3** In **Search**, enter **yarn.scheduler.maximum-allocation-mb** to modify the parameter, save the configuration, and then restart the service.
- End

## 19.8 Error "Can't get the Kerberos realm" Is Reported When a Spark Task Is Submitted in Yarn-Cluster Mode

### Symptom

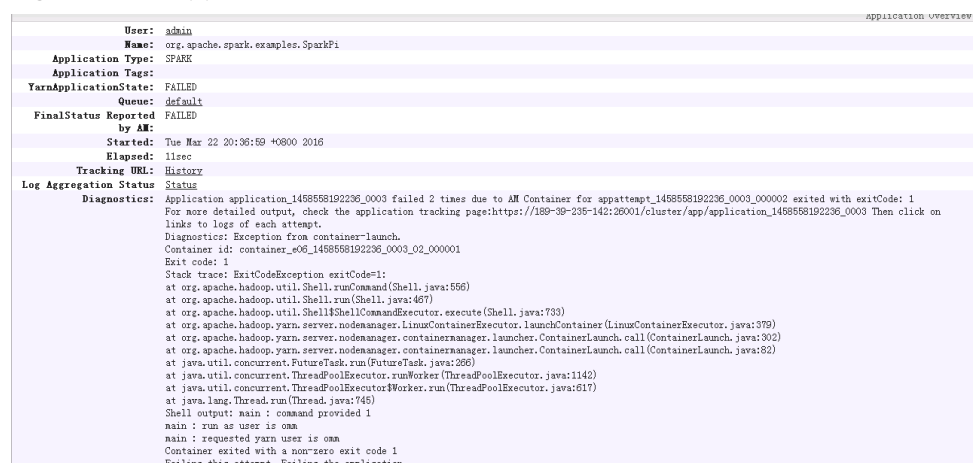
A Spark task fails to be submitted due to an authentication exception, and error message "Can't get the Kerberos realm" is reported.

### Cause Analysis

1. According to the exception printed in the driver log, the token information used to connect to HDFS cannot be found.
 

```
16/03/22 20:37:10 WARN Client: Exception encountered while connecting to the server :
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.token.SecretManager
$InvalidToken): token (HDFS_DELEGATION_TOKEN token 192 for admin) can't be found in cache
16/03/22 20:37:10 WARN Client: Failed to cleanup staging dir .sparkStaging/
application_1458558192236_0003
org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.token.SecretManager
$InvalidToken): token (HDFS_DELEGATION_TOKEN token 192 for admin) can't be found in cache
```
2. The native Yarn web UI shows that ApplicationMaster fails to be started twice and the task exits.

Figure 19-2 ApplicationMaster start failure



3. The ApplicationMaster log shows the following error information:
 

```
Exception in thread "main" java.lang.ExceptionInInitializerError
Caused by: org.apache.spark.SparkException: Unable to load YARN support
Caused by: java.lang.IllegalArgumentException: Can't get Kerberos realm
Caused by: java.lang.reflect.InvocationTargetException
```

```
Caused by: KrbException: Cannot locate default realm
Caused by: KrbException: Generic error (description in e-text) (60) - Unable to locate Kerberos realm
org.apache.hadoop.hive.metastore.MetaStoreUtils.newInstance(MetaStoreUtils.java:1410)
... 86 more
Caused by: javax.jdo.JDOFatalInternalException: Unexpected exception caught.
NestedThrowables:java.lang.reflect.InvocationTargetException
... 110 more
```

4. When you execute `./spark-submit --class yourclassname --master yarn-cluster /yourdependencyjars` to submit a task in Yarn-cluster mode, the driver is started in the cluster. Because the client's `spark.driver.extraJavaOptions` is loaded, you cannot find the `kdc.conf` file in the target path on the cluster node and cannot obtain information required for Kerberos authentication. As a result, the ApplicationMaster fails to be started.

## Solution

When submitting a task on the client, configure the `spark.driver.extraJavaOptions` parameter in the CLI. In this way, the `spark.driver.extraJavaOptions` parameter in the `spark-defaults.conf` file is not automatically loaded from the client path. When starting a Spark task, use `--conf` to specify the driver configuration as follows (note that the quotation mark after `spark.driver.extraJavaOptions=` is mandatory):

```
./spark-submit -class yourclassname --master yarn-cluster --conf
spark.driver.extraJavaOptions="
-Dlog4j.configuration=file:/opt/client/Spark/spark/conf/log4j.properties -
Djetty.version=x.y.z -Dzookeeper.server.principal=zookeeper/
hadoop.794bbab6_9505_44cc_8515_b4eddc84e6c1.com -
Djava.security.krb5.conf=/opt/client/KrbClient/kerberos/var/krb5kdc/
krb5.conf -Djava.security.auth.login.config=/opt/client/Spark/spark/conf/
jaas.conf -Dorg.xerial.snappy.tmpdir=/opt/client/Spark/tmp -
Dcarbon.properties.filepath=/opt/client/Spark/spark/conf/
carbon.properties" ../yourdependencyjars
```

## 19.9 Failed to Start spark-sql and spark-shell Due to JDK Version Mismatch

### Symptom

The JDK version does not match. As a result, the client fails to start spark-sql and spark-shell.

### Cause Analysis

1. The following error information is displayed on the Driver:  
Exception Occurs: BadPadding 16/02/22 14:25:38 ERROR Schema: Failed initialising database. Unable to open a test connection to the given database. JDBC url = `jdbc:postgresql://ip:port/sparkhivemeta`, username = spark. Terminating connection pool (set lazyInit to true if you expect to start your database after your app).
2. When a SparkSQL task is used, DBService needs to be accessed to obtain metadata information. On the client, the ciphertext needs to be decrypted for access. During the use, the user does not follow the process or configure environment variables, and the default JDK version exists in the environment

variables of the client. As a result, the decryption program invoked during decryption is abnormal, and the user is locked.

## Solution

**Step 1** Run the **which java** command to check whether the default Java command is the Java command of the client.

**Step 2** If it is not, go to the next step.

```
source ${client_path}/bigdata_env
```

Run the **kinit username** command and enter the password corresponding to the username to start the task.

----End

## 19.10 ApplicationMaster Fails to Start Twice When a Spark Task Is Submitted in Yarn-client Mode

### Symptom

In Yarn-client mode, ApplicationMaster fails to start twice.

### Cause Analysis

1. Driver exception:  
16/05/11 18:10:56 INFO Client:  
client token: N/A  
diagnostics: Application application\_1462441251516\_0024 failed 2 times due to AM Container for appattempt\_1462441251516\_0024\_000002 exited with exitCode: 10  
For more detailed output, check the application tracking page:[https://hdnode5:26001/cluster/app/application\\_1462441251516\\_0024](https://hdnode5:26001/cluster/app/application_1462441251516_0024) Then click on links to logs of each attempt.  
Diagnostics: Exception from container-launch.  
Container id: container\_1462441251516\_0024\_02\_000001
2. The ApplicationMaster log file contains the following error information:  
2016-05-12 10:21:23,715 | ERROR | [main] | Failed to connect to driver at 192.168.30.57:23867, retrying ... | org.apache.spark.Logging\$class.logError(Logging.scala:75)  
2016-05-12 10:21:24,817 | ERROR | [main] | Failed to connect to driver at 192.168.30.57:23867, retrying ... | org.apache.spark.Logging\$class.logError(Logging.scala:75)  
2016-05-12 10:21:24,918 | ERROR | [main] | Uncaught exception: | org.apache.spark.Logging\$class.logError(Logging.scala:96)  
org.apache.spark.SparkException: Failed to connect to driver!  
at org.apache.spark.deploy.yarn.ApplicationMaster.waitForSparkDriver(ApplicationMaster.scala:426)  
at org.apache.spark.deploy.yarn.ApplicationMaster.runExecutorLauncher(ApplicationMaster.scala:292)  
...  
2016-05-12 10:21:24,925 | INFO | [Thread-1] | Unregistering ApplicationMaster with FAILED (diag message: Uncaught exception: org.apache.spark.SparkException: Failed to connect to driver!) | org.apache.spark.Logging\$class.logInfo(Logging.scala:59)

In Spark-client mode, the task Driver runs on a client node (usually a node outside the cluster). During the startup, the ApplicationMaster process is started in the cluster. After the process is started, information needs to be registered with the Driver process. The task can be continued only after the registration is successful. According to the ApplicationMaster log, the connection to the Driver fails, which causes the task failure.

## Solution

**Step 1** Check whether the IP address of the Driver process can be pinged.

**Step 2** Start a Spark PI task. Information similar to the following is displayed:

```
16/05/11 18:07:20 INFO Remoting: Remoting started; listening on addresses :[akka.tcp://  
sparkDriver@192.168.1.100:23662]  
16/05/11 18:07:20 INFO Utils: Successfully started service 'sparkDriver' on port 23662.
```

**Step 3** Run the **netstat - anp | grep 23662** command on the node (192.168.1.100 in [Step 2](#)) to check whether the port is enabled. The following information indicates that the port is enabled.

```
tcp      0      0 ip:port :::*          LISTEN      107274/java  
tcp      0      0 ip:port ip:port      ESTABLISHED 107274/java
```

**Step 4** Run the **telnet 192.168.1.100 23662** command on the node where ApplicationMaster is started to check whether the port can be connected. Perform this operation as both the **root** and **omm** users. If information similar to **Escape character is '^[]'** is displayed, the connection is normal. If **connection refused** is displayed, the connection fails and the related port cannot be connected.

If the port is enabled but cannot be connected from other nodes, check the network configuration.

### NOTE

The port (port 23662 in this example) is randomly selected each time. Therefore, you need to test the port enabled by the task.

----End

## 19.11 Failed to Connect to ResourceManager When a Spark Task Is Submitted

### Symptom

The connection to ResourceManager is abnormal, and Spark tasks fail to be submitted.

### Cause Analysis

1. The following error message is displayed on the Driver, indicating that connections to port 26004 on the active and standby ResourceManager nodes are denied:

```
15/08/19 18:36:16 INFO RetryInvocationHandler: Exception while invoking getClusterMetrics of class  
ApplicationClientProtocolPBClientImpl over 33 after 1 fail over attempts. Trying to fail over after  
sleeping for 17448ms.  
java.net.ConnectException: Call From ip0 to ip1:26004 failed on connection exception:  
java.net.ConnectException: Connection refused.  
INFO RetryInvocationHandler: Exception while invoking getClusterMetrics of class  
ApplicationClientProtocolPBClientImpl over 32 after 2 fail over attempts. Trying to fail over after  
sleeping for 16233ms.  
java.net.ConnectException: Call From ip0 to ip2:26004 failed on connection exception:  
java.net.ConnectException: Connection refused;
```

2. On MRS Manager, check whether ResourceManager is running properly. If Yarn is faulty or an unknown exception occurs on a Yarn service instance, the ResourceManager of the cluster may be abnormal.

3. Check whether the client in the cluster is of the latest version.  
Check whether the ResourceManager instance has been migrated in the cluster. (Uninstall a ResourceManager instance and add it back to other nodes.)
4. On MRS Manager, view audit logs and check whether related operations are recorded.  
Run the **ping** command to check whether the IP address can be pinged.

## Solution

- If ResourceManager is abnormal, rectify the fault by referring to the related alarm handling section.
- If the client is not the latest, download and install the client again.
- If the IP address cannot be pinged, contact the network administrator to check the network.
- If HA is enabled for the cluster, set Yarn parameter **yarn.client.failover-sleep-base-ms** to a smaller value.

## 19.12 DataArts Studio Failed to Schedule Spark Jobs

### Symptom

DataArts Studio fails to schedule jobs, and the following error is reported indicating that data in the **/thriftserver/active\_thriftserver** directory cannot be read.

The error information is as follows:

```
Can not get JDBC Connection, due to KeeperErrorCode = NoNode for /thriftserver/active_thriftserver.
```

### Cause Analysis

When DataArts Studio submits a Spark job, Spark JDBC is invoked. Spark starts a ThriftServer process for the client to provide JDBC connections. During the startup, JDBCServer creates the **active\_thriftserver** subdirectory in the **/thriftserver** directory of ZooKeeper, and registers related connection information. If the connection information cannot be read, the JDBC connection is abnormal.

### Procedure

Check whether the ZooKeeper directory contains the target directory and registration information.

**Step 1** Log in to any master node as user **root** and initialize environment variables.

```
source /opt/client/bigdata_env
```

**Step 2** Run the **zkCli.sh -server 'ZooKeeper instance IP address:ZooKeeper connection port'** command to log in to ZooKeeper.

Generally, the ZooKeeper connection port number is 2181. You can obtain the port number from the ZooKeeper configuration parameter **clientPort**.

- Step 3** Run the `ls /thriftserver` command to check whether the `active_thriftserver` directory exists.
- If the `active_thriftserver` directory exists, run the `get /thriftserver/active_thriftserver` command to check whether it contains the registered configuration information.
    - If yes, contact Huawei Cloud technical support.
    - If no, go to [Step 4](#).
  - If the `active_thriftserver` directory does not exist, go to [Step 4](#).
- Step 4** Log in to Manager and check whether the active/standby status of the Spark JDBCServer instance is unknown.
- If yes, go to [Step 5](#).
  - If no, contact O&M personnel.
- Step 5** Restart the two JDBCServer instances. Check whether the status of the active and standby instances is normal and whether the target directory and data exist in ZooKeeper. If yes, the job is restored. If the instance status is not restored, contact Huawei Cloud technical support.
- End

## 19.13 Job Status Is error After a Spark Job Is Submitted Through an API

### Issue

After a Spark job is submitted using an API, the job status is displayed as **error**.

### Symptom

After the log level in `/opt/client/Spark/spark/conf/log4j.properties` is changed and a job is submitted using API V1.1, the job status is displayed as **error**.

### Cause Analysis

The executor monitors the job log output and determines the job execution result. After the execution result is changed to **error**, the output result cannot be detected. Therefore, the executor determines that the job status is abnormal after the job expires.

### Procedure

Change the log level in the `/opt/client/Spark/spark/conf/log4j.properties` file to **info**.

### Summary and Suggestions

You are advised to use the V2 API to submit jobs.



## 19.14 ALM-43006 Is Repeatedly Reported for the MRS Cluster

### Issue

Alarm **ALM-43006 Heap Memory Usage of the JobHistory Process Exceeds the Threshold** is repeatedly reported for the cluster, and handling according to the alarm reference is invalid.

### Symptom

Alarm **ALM-43006 Heap Memory Usage of the JobHistory Process Exceeds the Threshold** is generated for the cluster. The same alarm is generated again a period of time after handling measures are taken.

### Cause Analysis

The JobHistory memory leakage may occur. You need to install the corresponding patch to rectify the fault.

### Procedure

- Increase the heap memory of the JobHistory process.
- If the heap memory has been increased, restart the JobHistory instance.

## 19.15 Failed to Create or Delete a Table in Spark Beeline

### Issue

When the customer frequently creates or deletes a large number of users in Spark Beeline, some users occasionally fail to create or delete tables.

### Symptom

The procedure for creating a table is as follows:

```
CREATE TABLE wlg_test001 (start_time STRING,value INT);
```

The following error message is displayed:

```
Error: org.apache.spark.sql.AnalysisException:  
org.apache.hadoop.hive.ql.metadata.HiveException: MetaException(message:Failed to grant permission on  
HDFSjava.lang.reflect.UndeclaredThrowableException); (state=,code=0)
```

### Cause Analysis

1. View MetaStore logs.

```

hive.metastore.RetryingHMSHandler: | org.apache.hadoop.hive.ql.log.PerfLogger.PerfLogBegin(PerfLogger.java:121)
2020-08-31 14:41:38,504 | INFO | pool-7-thread-197 | 197: create_table: Table(tableName:wlg_test001, dbName:hive_csb_csb_3f8_x48s
srbt_5lbi2edu, owner:CSB_csb_3f8_x48ssrbt, createTime:1598856098, lastAccessTime:0, retention:0, sd:StorageDescriptor(cols:[FieldS
chema(name:start_time, type:string, comment:null), FieldSchema(name:value, type:int, comment:null)], location:hdfs://hacluster/use
r/hive/warehouse/hive_csb_csb_3f8_x48ssrbt_5lbi2edu.db/wlg_test001, inputFormat:org.apache.hadoop.mapred.TextInputFormat, outputFo
rmat:org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat, compressed:false, numBuckets:-1, serDeInfo:SerDeInfo(name:null, s
erializationLib:org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe, parameters:{serialization.format=1}), bucketCols:[], sortCols:
[], parameters:{}, skewedInfo:SkewedInfo(skewedColNames:[], skewedColValues:[], skewedColValueLocationMaps:{}), partitionKeys:[],
parameters:{spark.sql.sources.schema.numParts=1, spark.sql.sources.schema.part.0={"type":"struct","fields":{"name":"start_time",
"type":"string","nullable":true,"metadata":{}},"name":"value","type":"integer","nullable":true,"metadata":{}}}}, viewOriginalTex
t:null, viewExpandedText:null, tableType:MANAGED_TABLE, privileges:PrincipalPrivilegeSet(userPrivileges:{CSB_csb_3f8_x48ssrbt=[Pri
vilegeGrantInfo(privilege:INSERT, createTime:-1, grantor:spark, grantorType:USER, grantOption:true), PrivilegeGrantInfo(privilege:
SELECT, createTime:-1, grantor:spark, grantorType:USER, grantOption:true), PrivilegeGrantInfo(privilege:UPDATE, createTime:-1, gra
ntor:spark, grantorType:USER, grantOption:true), PrivilegeGrantInfo(privilege:DELETE, createTime:-1, grantor:spark, grantorType:US
ER, grantOption:true)]}, groupPrivileges:null, rolePrivileges:null) | org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.l
ogInfo(HiveMetaStore.java:881)
2020-08-31 14:41:38,515 | WARN | pool-7-thread-197 | Location: hdfs://hacluster/user/hive/warehouse/hive_csb_csb_3f8_x48ssrbt_5lbi
i2edu.db/wlg_test001 specified for non-external table:wlg_test001 | org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.crea
te_table_core(HiveMetaStore.java:1546)
2020-08-31 14:41:38,516 | INFO | pool-7-thread-197 | Creating directory if it doesn't exist: hdfs://hacluster/user/hive/warehouse
/hive_csb_csb_3f8_x48ssrbt_5lbi2edu.db/wlg_test001 | org.apache.hadoop.hive.common.FileUtils.mkdir(FileUtils.java:507)
2020-08-31 14:41:38,566 | INFO | pool-7-thread-197 | 197: get_database: hive_csb_csb_3f8_x48ssrbt_5lbi2edu | org.apache.hadoop.hi
ve.metastore.HiveMetaStore$HMSHandler.logInfo(HiveMetaStore.java:881)
2020-08-31 14:41:38,578 | INFO | pool-7-thread-197 | 197: get_table : db=hive_csb_csb_3f8_x48ssrbt_5lbi2edu tbl=wlg_test001 | org
.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.logInfo(HiveMetaStore.java:881)
2020-08-31 14:41:38,594 | ERROR | pool-7-thread-197 | MetaException(message:Failed to grant permission on HDFS)java.lang.reflect.Un
declaredThrowableException)
    at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.create_table_with_environment_context(HiveMetaStore.java:1638
)
    at sun.reflect.GeneratedMethodAccessor94.invoke(Unknown Source)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:498)
    at org.apache.hadoop.hive.metastore.RetryingHMSHandler.invokeInternal(RetryingHMSHandler.java:140)

```

2. View HDFS logs.

```

2020-08-31 14:41:38,568 | INFO | Socket Reader #1 for port 9820 | Authorization successful for hive/hadoop.036a3461_d09b_494f_a32
c_af273307d943.com@036A3461_D09B_494F_A32C_AF273307D943.COM (auth:KERBEROS) for protocol=interface org.apache.hadoop.hdfs.protocol
_ClientProtocol | ServiceAuthorizationManager.java:135
2020-08-31 14:41:38,586 | INFO | IPC Server handler 7 on 9820 | IPC Server handler 7 on 9820, call Call#3822197 Retry#0 org.apach
e.hadoop.hdfs.protocol_ClientProtocol.checkAccess from 192.168.1.66:50540: org.apache.hadoop.security.AccessControlException: Perm
ission denied: user=hive, access=READ, inode="/user/hive/warehouse/hive_csb_csb_3f8_x48ssrbt_5lbi2edu.db/wlg_test001":spark:hive:d
rwx----- | Server.java:2523
2020-08-31 14:41:38,852 | INFO | Socket Reader #1 for port 9820 | Auth successful for hwstaff_pub_0tw00ru6@036A3461_D09B_494F_A32
C_AF273307D943.COM (auth:TOKEN) | Server.java:1700
2020-08-31 14:41:38,911 | INFO | Socket Reader #1 for port 9820 | Authorization successful for hwstaff_pub_0tw00ru6@036A3461_D09B

```

3. Compare permissions (**test001** is a table created by a user in the abnormal state, and **test002** is a table created by a user in the normal state).

```

drwx----- - spark             hive             0 2020-08-31 14:41 /user/hive/warehouse/hive_csb_csb_3f8_x48ssrbt_5lbi2edu.db/wl
g_test001
drwxrwx---+ - spark             hive             0 2020-08-31 15:07 /user/hive/warehouse/hive_csb_csb_3f8_x48ssrbt_5lbi2edu.db/wl
g_test002

```

4. An error similar to the following is reported when a table is dropped:

```

dataplan_modela_csbch2;
Error: Error while compiling statement: FAILED:
SemanticException Unable to fetch table dataplan_modela_csbch2.
java.security.AccessControlException: Permission denied: user=CSB_csb_3f8_x48ssrbt,
access=READ,
inode="/user/hive/warehouse/hive_csb_csb_3f8_x48ssrbt_5lbi2edu.db/
dataplan_modela_csbch2":spark:hive:drwx-----

```

5. Analyze the cause.

The default user created during cluster creation uses the same UID, causing user disorder. This problem is triggered when a large number of users are created. As a result, the Hive user does not have the permission to create tables occasionally.

```

[root@node-master21Mrt ~]#
[root@node-master21Mrt ~]#
[root@node-master21Mrt ~]# id hive
uid=20013(hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com) gid=10002(hive) groups=10002(hive)
[root@node-master21Mrt ~]#
[root@node-master21Mrt ~]#
[root@node-master21Mrt ~]#
[root@node-master21Mrt ~]# id hive
uid=20013(hive) gid=10002(hive) groups=10002(hive),10001(hadoop),10000(supergroup),8003(System_administrator_186),9998(ficommon)
[root@node-master21Mrt ~]#
[root@node-master21Mrt ~]#

```

```
objectClass: krbPrincipalAux
objectClass: krbTicketPolicyAux
# hive, Peoples, hadoop.com
dn: cn=hive,ou=Peoples,dc=hadoop,dc=com
uid: hive
homeDirectory: /home/hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com
cn: hive
uidNumber: 20013
objectClass: account
objectClass: posixAccount
objectClass: shadowAccount
userPassword:: e1NTSEF9cXZWS0VlMi9pYVFpdzFmUmNIUVJFUEJYZWtKlZHMhk=
gidNumber: 10002
# hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com, Peoples, hadoop.com
dn: cn=hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com,ou=Peoples,dc=hadoop,dc=com
uid: hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com
homeDirectory: /home/hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com
cn: hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com
uidNumber: 20013
objectClass: account
objectClass: posixAccount
objectClass: shadowAccount
gidNumber: 10002
description: [userName:"hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com"]
description: [userType:"1"]
description: [groupList:"hive,hadoop,supergroup,compcommon"]
description: [roleList:"System administrator"]
description: [description:"aGL2ZS8kZWZhdWw0IHVzZXIjSGl2Zem7m0iup0eUq0aItw=="]
description: [createTime:"1554974652422"]
description: [defaultUser:"0"]
description: [primaryGroup:"hive"]
# hive/hadoop.036a3461_d09b_494f_a32c_af273307d943.com@036A3461_D09B_494F_A32C_AF273307D943.COM, 036A3461_D09B_494F_A32C_AF273307D943.COM, krbcontainer, hado
```

## Procedure

Restart the `sssd` process of the cluster.

Run the `service sssd restart` command as user `root` to restart the `sssd` process and run the `ps -ef | grep sssd` command to check whether the `sssd` process is running properly.

In normal cases, the `/usr/sbin/sssd` process and three sub-processes `/usr/libexec/sssd/sss_dbe`, `/usr/libexec/sssd/sss_nss` and `/usr/libexec/sssd/sss_pam` exist.

# 19.16 Failed to Connect to the Driver When a Spark Job Is Submitted on a Node Outside the Cluster

## Symptom

Nodes outside the cluster can communicate with each node in the cluster. When a node outside the cluster submits a Spark task to Yarn in client mode, the submission fails and an error message is displayed, indicating that the Driver cannot be connected.

## Cause Analysis

When a Spark task is submitted in the client mode, the driver process of Spark is on the client side, and the executor needs to interact with the driver to run the job.

If the NodeManager fails to connect to the node where the client is located, the following error is reported:

```
Log Length: 174453
Showing 4096 bytes of 174453 total. Click here for the full log.
connect to driver at ecs-d6d9-1112169:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,150 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,251 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,351 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,452 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,552 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,653 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,753 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,855 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:34,956 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:35,057 ERROR | [main] | Failed to connect to driver at [redacted]:22741, retrying ... | org.apache.spark.internal.Logging$class.logError(Logging.scala:70)
2020-11-21 16:04:35,161 ERROR | [main] | Uncaught exception: | org.apache.spark.internal.Logging$class.logError(Logging.scala:91)
org.apache.spark.SparkException: Failed to connect to driver!
    at org.apache.spark.deploy.yarn.ApplicationMaster.waitForSparkDriver(ApplicationMaster.scala:630)
```

## Procedure

Specify the IP address of the driver in the Spark configuration of the client.

Add **spark.driver.host=driverIP** to *<Client installation path>/Spark/spark/conf/spark-defaults.conf* and run the Spark task again.

## Summary and Suggestions

You are advised to submit jobs in cluster mode.

# 19.17 Large Number of Shuffle Results Are Lost During Spark Task Execution

## Symptom

Spark tasks fail to be executed. The task log shows that a large number of **shuffle** files are lost.

## Cause Analysis

When Spark is running, the **shuffle** file generated temporarily is stored in the temporary directory of the executor for later use.

When an executor exits abnormally, NodeManager deletes the temporary directory of the container where the executor is located. When other executors apply for the shuffle result of the executor, a message is displayed indicating that the file cannot be found.

Therefore, you need to check whether the executor exits abnormally. You can check whether there are executors in the **dead** state on the executors tab page on the Spark task page and view the executor logs of each **dead** state, determine the cause of abnormal exit. Some executors may exit because the **shuffle** file cannot be found. You need to find the earliest executor that exits abnormally.

Common abnormal exit causes:

- OOM occurs on the executor.
- Multiple tasks fail when the executor is running.
- The node where the executor is located is cleared.

## Procedure

Adjust or modify the task parameters or code based on the actual cause of the abnormal exit of the executor, and run the Spark task again.

## 19.18 Disk Space Is Insufficient Due to Long-Term Running of JDBCServer

### Issue

When the JDBCServer service connected to Spark submits a spark-sql task to the Yarn cluster, the data disk of the Core node is fully occupied after the task runs for a period of time.

### Symptom

When the JDBCServer service of a customer connected to Spark submits a spark-sql task to the Yarn cluster, the data disk of the Core node is fully occupied after the task runs for a period of time.

After checking the disk usage in the background, it is found that there are too many APP temporary files (files generated by shuffle) of the JDBCServer service, and the files are not cleared, occupying a large amount of memory.

### Cause Analysis

After checking the directories that contain a large number of files on the Core node, it is found that most of the directories are similar to **blockmgr-033707b6-fbbb-45b4-8e3a-128c9bcfa4bf**, which stores temporary shuffle files generated during computing.

The dynamic resource allocation function of Spark is enabled on JDBCServer, and shuffle is hosted by NodeManager. NodeManager only manages these files based on the running period of the application, and does not check whether the container where a single executor is located exists. Therefore, the temporary files are deleted only when the app is stopped. When a task runs for a long time, a large number of temporary files occupy a large amount of disk space.

### Procedure

Start a scheduled task to delete shuffle files that have been stored for a specified period of time. For example, delete shuffle files that have been stored for more than 6 hours each hour.

**Step 1** Create the **clean\_appcache.sh** script. If there are multiple data disks, change the value of **data1** in **BASE\_LOC** based on the actual situation.

- Security cluster

```
#!/bin/bash
BASE_LOC=/srv/BigData/hadoop/data1/nm/localdir/usercache/spark/appcache/application_*/
blockmgr*
find $BASE_LOC/ -mmin +360 -exec rmdir {} \;
find $BASE_LOC/ -mmin +360 -exec rm {} \;
```

- Common cluster

```
#!/bin/bash
BASE_LOC=/srv/BigData/hadoop/data1/nm/localdir/usercache/omm/appcache/application_*/
blockmgr*
find $BASE_LOC/ -mmin +360 -exec rmdir {} \;
find $BASE_LOC/ -mmin +360 -exec rm {} \;
```

**Step 2** Run the following commands to change the permission to the script:

```
chmod 755 clean_appcache.sh
```

**Step 3** Add a scheduled task to start the clearance script. Change the script path to the actual path.

Run the **crontab -l** command to view the scheduled task.

Run the **crontab -e** command to edit the scheduled task.

```
0 * * * * sh /root/clean_appcache.sh > /dev/null 2>&1
```

```
----End
```

## 19.19 Failed to Load Data to a Hive Table Across File Systems by Running SQL Statements Using Spark Shell

### Issue

When the **spark-shell** command is used to execute SQL statements or the **spark-submit** command is used to submit Spark tasks, the **load** command of SQL statements exists, and the source data and target table are not stored in the same file system. An error is reported when the MapReduce task is started in the preceding two modes.

### Cause Analysis

When the **load** command is used to import data to the Hive table across file systems (for example, the original data is stored in HDFS but the Hive table data is stored in OBS), and the file length is greater than the threshold (32 MB by default). In this case, the MapReduce task that uses DistCp is triggered to migrate data. The MapReduce task configuration is directly extracted from the Spark task configuration. However, the **net.topology.node.switch.mapping.impl** configuration item of the Spark task does not retain the default value of Hadoop. Therefore, the JAR package of Spark needs to be used. As a result, MapReduce reports an error indicating that the class cannot be found.

### Procedure

Solution 1:

If the file size is small, set the default file size to a value greater than the maximum file size. For example, if the maximum file size is 95 MB, run the following command:

```
hive.exec.copyfile.maxsize=104857600
```

Solution 2:

If the file size is large, use DistCp to improve the data migration efficiency. Add the following parameters when starting the Spark task:

```
--conf spark.hadoop.net.topology.node.switch.mapping.impl=org.apache.hadoop.net.ScriptBasedMapping
```



## 19.21 Spark Task Execution Failure

### Symptom

- An executor out of memory (OOM) error occurs.
- The information about the failed task shows that the failure cause is "lost task xxx."

### Cause Analysis

- Symptom 1: The data volume is too large or too many tasks are running on the same executor at the same time.
- Symptom 2: Some tasks fail to be executed. When the error is reported, determine the node where the lost task is running. Generally, the error is caused by the abnormal exit of the lost task.

### Procedure

- Symptom 1:
  - If the data volume is too large, adjust the memory size of the executor and use **--executor-memory** to specify the memory size.
  - If too many tasks are running at the same time, check the number of vcores specified by **--executor-cores**.
- Symptom 2:

Locate the cause in the corresponding task log. If an OOM error occurs, see the solution to symptom 1.

## 19.22 JDBCServer Connection Failure

### Symptom

- The ha-cluster cannot be identified (unknowHost or port required).
- Failed to connect to JDBCServer.

### Cause Analysis

- Symptom 1: The **spark-beeline** command is used to connect to JDBCServer. JDBCServer in versions earlier than MRS 3.0 adopts HA. Therefore, a specific URL and the JAR package provided by MRS Spark are required to connect to JDBCServer.
- Symptom 2: The JDBCServer service is not running properly or port monitoring is abnormal.

### Procedure

- Symptom 1: Use a specific URL and the JAR package provided by MRS Spark to connect to JDBCServer.
- Symptom 2: Check that the JDBCServer service is running properly and port monitoring is normal, and try again.



## 19.23 Failed to View Spark Task Logs

### Symptom

- A user fails to view logs when a task is running.
- A user fails to view logs when a task is complete.

### Cause Analysis

- Symptom 1: The MapReduce component is abnormal.
- Symptom 2:
  - The JobHistory service of Spark is abnormal.
  - The log size is too large, and NodeManager times out during log aggregation.
  - The permission on the HDFS log storage directory (`/tmp/logs/Username/logs` by default) is abnormal.
  - Logs have been deleted. By default, Spark JobHistory stores event logs for seven days (specified by `spark.history.fs.cleaner.maxAge`). MapReduce stores task logs for 15 days (specified by `mapreduce.jobhistory.max-age-ms`).
  - If the task cannot be found on the Yarn page, it may have been cleared by Yarn. By default, Yarn stores 10,000 historical tasks (specified by `yarn.resourcemanager.max-completed-applications`).

### Procedure

- Symptom 1: Check whether the MapReduce component is running properly. If it is abnormal, restart it. If the fault persists, check the JobHistoryServer log file in the background.
- Symptom 2: Perform the following checks in sequence:
  - a. Check whether JobHistory of Spark is running properly.
  - b. On the app details page of Yarn, check whether the log file is too large. If log aggregation fails, the value of **Log Aggregation Status** should be **Failed** or **Timeout**.
  - c. Check whether the permission on the corresponding directory is normal.
  - d. Check whether the corresponding **appid** file exists in the directory. In MRS 3.x or later, the event log files are stored in the `hdfs://hacluster/spark2xJobHistory2x` directory. In versions earlier than MRS 3.x, the event log files are stored in the `hdfs://hacluster/sparkJobHistory` directory. The task run logs are stored in the `hdfs://hacluster/tmp/logs/Username/logs` directory.
  - e. Check whether **appid** or the current job ID exceeds the maximum value in the historical records.

## 19.24 Spark Streaming Task Submission Issues

### Symptom

- A message is displayed indicating that the class cannot be found when the Kafka is connected.
- An authentication error is reported when Kafka with Kerberos is connected.
- After the SparkStreaming task runs for a period, a message is displayed indicating that the token has expired.

### Cause Analysis

- Symptom 1: By default, the Kafka JAR package is not loaded when Spark submits tasks. Therefore, **--jars** needs to be added to the startup command to specify the JAR package corresponding to the Kafka version.
- Symptom 2: The Spark authentication information cannot be used for connecting to Kafka. JVM parameters required for authentication must be set.
- Symptom 3: By default, Spark uses the authentication information of the current client to submit tasks. The code login mode can also be used. However, neither of the two authentication modes can update the token used by the task. When the token information generated during submission expires, the token cannot be used. As a result, an error is reported.

### Procedure

- Symptom 1: Add **--jars** to the startup command to specify the JAR package corresponding to the Kafka version. Generally, the JAR package is in the Spark client directory **/jars/streamingClient** (Kafka 0.8) or **/jars/streamingClient010** (Kafka 0.10).
- Symptom 2: Compile and run the application. For details, see [Compiling and Running Applications](#).
- Symptom 3: Use **--keytab** and **--principal** to add the keytab file and the corresponding user to the task. If the keytab file is the same as that configured in **jaas.conf** of Kafka, Spark reports an error indicating that a file is uploaded for multiple times. The solution is to copy a keytab file so that different files are uploaded by **--files** and **--keytab**.

## 19.25 Authentication Fails When Spark Connects to Other Services

### Symptom

- When Spark connects to HBase, an authentication failure message is displayed or the HBase table cannot be connected.
- When Spark connects to HBase, a message is displayed indicating that the JAR package cannot be found.

## Cause Analysis

- Symptom 1: HBase does not obtain the authentication information of the current task. As a result, the authentication fails when HBase is connected, and the corresponding data cannot be read.
- Symptom 2: By default, Spark does not load the HBase JAR package. You need to use `--jars` to add the JAR package to the task.

## Procedure

- Symptom 1: Enable the HBase authentication function by running the `spark.yarn.security.credentials.hbase.enabled=true` command. However, do not replace `hbase-site.xml` on the Spark client with `hbase-site.xml` on the HBase client because they are not completely consistent.
- Symptom 2: Use `--jars` to upload the HBase JAR package.

# 19.26 Authentication Fails When Spark Connects to Kafka

## Symptom

In an MRS 2.1.0 cluster, when Spark consumes Kafka data, an authentication failure occurs one day later.

```
2021-06-23 10:36:00,218 | INFO | Driver | Final app status: FAILED, exitCode: 15, (reason: User class threw exception: org.apache.spark.SparkException: Job aborted due to stage failure: Task 1 in stage 1515.0 failed 4 times, most recent failure: Lost task 1:3 in stage 1515.0 (TID 3036, node-ana-00e100b, executor 1): org.apache.kafka.common.errors.SaslAuthenticationException: An error: (java.security.PrivilegedActionException: java.security.sasl.SaslException: GSS initiate failed [Caused by GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt)]) occurred when evaluating SASL token received from the Kafka Broker. Kafka Client will go to AUTHENTICATION_FAILED state.
Caused by: java.security.sasl.SaslException: GSS initiate failed [Caused by GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt)]
at com.sun.security.sasl.gsskerb.GssKrb5Client.evaluateChallenge(GssKrb5Client.java:211)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator$2.run(SaslClientAuthenticator.java:362)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator$2.run(SaslClientAuthenticator.java:360)
at java.security.AccessController.doPrivileged(Native Method)
at java.security.auth.Subject.doAs(Subject.java:422)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator.createSaslToken(SaslClientAuthenticator.java:360)
at org.apache.kafka.common.security.authenticator.SaslClientAuthenticator.sendSaslClientToken(SaslClientAuthenticator.java:270)
at org.apache.kafka.common.network.KafkaChannel.prepare(KafkaChannel.java:83)
at org.apache.kafka.common.network.Selector.pollSelectionKeys(Selector.java:486)
at org.apache.kafka.common.network.Selector.poll(Selector.java:424)
at org.apache.kafka.clients.NetworkClient.poll(NetworkClient.java:460)
at org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.poll(ConsumerNetworkClient.java:261)
at org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.poll(ConsumerNetworkClient.java:233)
at org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.poll(ConsumerNetworkClient.java:224)
at org.apache.kafka.clients.consumer.internals.ConsumerNetworkClient.awaitMetadataUpdate(ConsumerNetworkClient.java:156)
at org.apache.kafka.clients.consumer.internals.ConsumerCoordinator.poll(ConsumerCoordinator.java:304)
at org.apache.kafka.clients.consumer.KafkaConsumer.pollOnce(KafkaConsumer.java:1149)
at org.apache.kafka.clients.consumer.KafkaConsumer.poll(KafkaConsumer.java:1115)
at org.apache.spark.streaming.kafka010.CacheKafkaConsumer.poll(CacheKafkaConsumer.scala:154)
at org.apache.spark.streaming.kafka010.CachedKafkaConsumer.get(CachedKafkaConsumer.scala:86)
at org.apache.spark.streaming.kafka010.KafkaRDDIterator.next(KafkaRDD.scala:279)
at org.apache.spark.streaming.kafka010.KafkaRDDIterator.next(KafkaRDD.scala:231)
```

## Cause Analysis

The cluster version does not match the version of the running program package.

## Procedure

1. Complete the configuration by referring to [Running the "Connecting Spark Streaming to Kafka0-10" Sample Application](#).
2. Obtain the `spark-streaming-kafka-0-10_2.11-2.3.2-mrs-2.1.jar` package if you need to use the sample programs provided by Maven, because the current cluster version is MRS 2.1.0.

## 19.27 An Error Occurs When SparkSQL Reads the ORC Table

### Symptom

An error occurs when SparkSQL is used to read an ORC table created by Hive.

```
Caused by: java.util.concurrent.ExecutionException: java.lang.IndexOutOfBoundsException: Index: 0
    at java.util.concurrent.FutureTask.report(FutureTask.java:122)
    at java.util.concurrent.FutureTask.get(FutureTask.java:192)
    at org.apache.hadoop.hive.q1.io.orc.OrcInputFormat.generateSplitsInfo(OrcInputFormat.java:1016)
    ... 51 more
Caused by: java.lang.IndexOutOfBoundsException: Index: 0
    at java.util.Collections$EmptyList.get(Collections.java:4456)
    at org.apache.hadoop.hive.q1.io.orc.OrcProto$Type.getSubtypes(OrcProto.java:12240)
    at org.apache.hadoop.hive.q1.io.orc.ReaderImpl.getColumnIndicesFromNames(ReaderImpl.java:651)
    at org.apache.hadoop.hive.q1.io.orc.ReaderImpl.getRawDataSizeOfColumns(ReaderImpl.java:634)
    at org.apache.hadoop.hive.q1.io.orc.OrcInputFormat$SplitGenerator.populateAndCacheStripeDetails(OrcInputFormat.java:927)
    at org.apache.hadoop.hive.q1.io.orc.OrcInputFormat$SplitGenerator.call(OrcInputFormat.java:836)
    at org.apache.hadoop.hive.q1.io.orc.OrcInputFormat$SplitGenerator.call(OrcInputFormat.java:702)
    at java.util.concurrent.FutureTask.run(FutureTask.java:266)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:748)
```

### Cause Analysis

This problem is caused by an open-source issue which is described at <https://issues.apache.org/jira/browse/HIVE-11102>. It may be triggered when the open-source Hive 1.2.1 package is used.

This problem has been resolved in the Hive for Spark module of MRS.

### Procedure

Use the Hive for Spark packages provided by the MRS cluster.

```
hive-beeline-1.2.1.spark_2.2.1-mrs-x.x.x.jar
hive-cli-1.2.1.spark_2.2.1-mrs-x.x.x.jar
hive-common-1.2.1.spark_2.2.1-mrs-x.x.x.jar
hive-exec-1.2.1.spark_2.2.1-mrs-x.x.x.jar
hive-jdbc-1.2.1.spark_2.2.1-mrs-x.x.x.jar
hive-metastore-1.2.1.spark_2.2.1-mrs-x.x.x.jar
```

Obtain the sample project from the Maven repository by referring to [Obtaining Sample Projects from Huawei Mirrors](#).

## 19.28 Failed to Switch to the Log Page from stderr and stdout on the Native Spark Web UI

### Symptom

When Spark tasks are executed in an MRS 3.x cluster, stderr and stdout on the native Spark2x web UI cannot redirect to the log page.

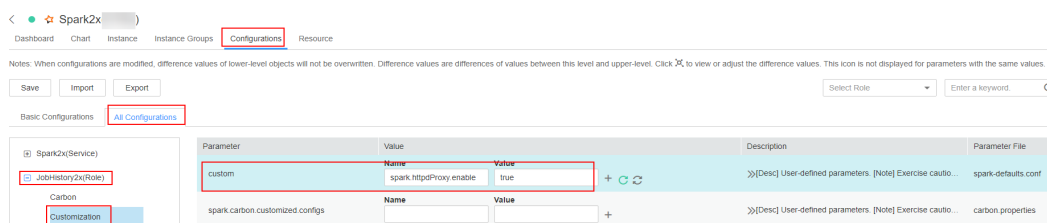
### Cause Analysis

The `spark.httpdProxy.enable` configuration item is not enabled.

## Procedure (Modifying Configuration Items)

- Step 1** On the client, set `spark.httpdProxy.enable` of the `spark-defaults.conf` configuration item to `true`. The default value is `false`.
- Step 2** Log in to Manager, choose **Cluster > Services > Spark2x**. On the Spark2x page, choose **Configurations > All Configurations > JobHistory2x > Customization**. Locate the `custom` parameter in the right pane, add configuration item `spark.httpdProxy.enable`, and set the value to `true`.

**Figure 19-3** Customizing a configuration item



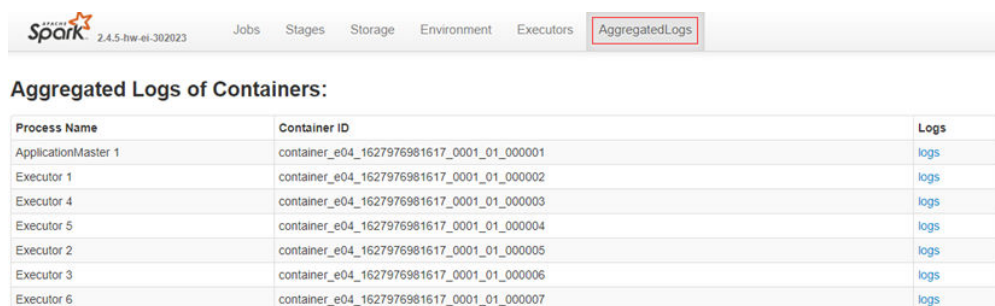
- Step 3** By referring to step 2, add configuration item `spark.httpdProxy.enable` and set the value to `true` respectively for `JDBCServer2x` and `SparkResource2x`.
- Step 4** Save the configuration. In the upper right corner of the Spark2x page, click **More** and select **Service Rolling Restart**.
- Step 5** Log in to the native Spark2x page and click `stdout` or `stderr` to access the executor log page.

----End

## Procedure (Without Modifying Configuration Items)

- View the logs of completed tasks.  
On the native Spark2x page, click **AggregatedLogs** to view the aggregated task logs.

**Figure 19-4** AggregatedLogs



- View the logs of running tasks.  
On the native Spark2x page, choose **Jobs > Stage > Task** to locate the target executor and then click `stdout` or `stderr` to view logs of the executor.

Figure 19-5 Jobs

The screenshot shows the 'Jobs' tab in the Spark UI. The page title is 'Spark Jobs (?)'. Below the title, it displays 'User: root', 'Total Uptime: 5.2 min', 'Scheduling Mode: FIFO', and 'Completed Jobs: 1'. There are links for 'Event Timeline' and 'Completed Jobs (1)'. A table lists the job details:

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total
0	count at <console>-25 count at <console>-25	2021/08/03 16:17:31	3 s	1/1

Figure 19-6 Stage

The screenshot shows the 'Stages' tab for Job 0. The page title is 'Details for Job 0'. It displays 'Status: SUCCEEDED', 'Completed Stages: 1', and links for 'Event Timeline', 'DAG Visualization', and 'Completed Stages (1)'. A table lists the stage details:

Stage Id	Description	Submitted	Duration	Tasks: Succeeded
0	count at <console>-25	2021/08/03 16:17:31	3 s	

Figure 19-7 Task

The screenshot shows the 'Tasks' tab for Stage 0. The page title is 'Details for Stage 0 (Attempt 0)'. It displays 'Total Time Across All Tasks: 4 s' and 'Locality Level Summary: Process local: 4'. There are links for 'DAG Visualization', 'Show Additional Metrics', and 'Event Timeline'. A table shows 'Summary Metrics for 4 Completed Tasks':

Metric	Min	25th percentile	Median	75th percentile	Max
Duration	0.8 s	0.9 s	1 s	1 s	1 s
GC Time	0 ms	0 ms	0 ms	13 ms	13 ms

Below this is a table for 'Aggregated Metrics by Executor':

Executor ID	Address	Task Time	Total Tasks	Failed Tasks	Killed Tasks	Succeeded Tasks	Blacklisted	Logs
1	node-ana-coreZHEL0003.8f5c96c-e74d-43f4-bafc-ab394f191507.com:22846	2 s	2	0	0	2	false	stdout stderr
2	node-ana-coreZHEL0002.8f5c96c-e74d-43f4-bafc-ab394f191507.com:22856	3 s	2	0	0	2	false	stdout stderr

At the bottom, there is a table for 'Tasks (4)':

Index	ID	Attempt	Status	Locality Level	Executor ID	Host	Launch Time	Duration	GC Time	Errors
0	0	0	SUCCESS	PROCESS_LOCAL	1	node-ana-coreZHEL0003.8f5c96c-e74d-43f4-bafc-ab394f191507.com	2021/08/03 16:17:32	0.8 s		
1	1	0	SUCCESS	PROCESS_LOCAL	2	node-ana-coreZHEL0002.8f5c96c-e74d-43f4-bafc-ab394f191507.com	2021/08/03 16:17:32	0.9 s	13 ms	
2	2	0	SUCCESS	PROCESS_LOCAL	1	node-ana-coreZHEL0003.8f5c96c-e74d-43f4-bafc-ab394f191507.com	2021/08/03 16:17:33	1 s		
3	3	0	SUCCESS	PROCESS_LOCAL	2	node-ana-coreZHEL0002.8f5c96c-e74d-43f4-bafc-ab394f191507.com	2021/08/03 16:17:33	1 s		

## Related Information

After you click **stdout** or **stderr** to go to the log page, only **stdout** or **stderr** logs are displayed. To view complete logs, you can delete **stdout** or **stderr** from the URL and then access all logs of the executor.

For example:

[https://<EIP>:9022/component/Yarn/NodeManager/15/node/containerlogs/container\\_e04\\_1627976981617\\_0002\\_01\\_000002/root/stderr?start=-4096](https://<EIP>:9022/component/Yarn/NodeManager/15/node/containerlogs/container_e04_1627976981617_0002_01_000002/root/stderr?start=-4096)

You can modify the URL as follows:

[https://<EIP>:9022/component/Yarn/NodeManager/15/node/containerlogs/container\\_e04\\_1627976981617\\_0002\\_01\\_000002/root/?start=-4096](https://<EIP>:9022/component/Yarn/NodeManager/15/node/containerlogs/container_e04_1627976981617_0002_01_000002/root/?start=-4096)









```

root@node-master2001:~# sqoop list-databases --connect jdbc:mysql://192.168.100.1:3306/ --username root --password root
Warning: /opt/hadoopclient/Sqoop/sqoop/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HDFS/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/Bigdata/client/HDFS/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/catalog/lib/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/geomesa/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/ranger-2.0.0-hbase-plugin/install/lib/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerB
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/lib/client-facing-thirdparty/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/lib/jdbc/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/tools/hbase-hbck2-...jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoopclient/HBase/HBase/tools/hbase-tools-...jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple-bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2022-01-29 10:56:53.892 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
2022-01-29 10:56:53.926 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
Sat Jan 29 10:56:54 CST 2022 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.20+
ents SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServer
--ssl-to-false-- You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verifi
information_schema
mysql
performance_schema
sys
test
test

```

The command output shows that Sqoop is successfully connected to the MySQL database.

----End

## 20.2 Failed to Find the HBaseAdmin.<init> Method When Sqoop Reads Data from the MySQL Database to HBase

### Issue

If the Sqoop client (version 1.4.7) of MRS is used to extract data from a specified table in the MySQL database to a table in HBase 2.2.3, the following exception is reported:

```

Trying to load data into HBASE through Sqoop getting below error.
Exception in thread "main" java.lang.NoSuchMethodError:
org.apache.hadoop.hbase.client.HBaseAdmin.<init>(Lorg/apache/hadoop/conf/Configuration;V

```

The following figure shows the complete exception information.

```

and provide truststore for server certificate verification.
2022-01-28 14:37:35.764 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `t_o_eso_users` AS t LIM
IT 1
2022-01-28 14:37:35.786 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `t_o_eso_users` AS t LIM
IT 1
2022-01-28 14:37:35.797 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/Bigdata/client/HDFS/hadoop
Note: /tmp/sqoop-root/compile/792dbda207bec0305d1989403855dfa2/t_o_eso_users.java uses or overrides a deprecated A
PI.
Note: Recompile with -Xlint:deprecation for details.
2022-01-28 14:37:36.678 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-root/compile/792dbda207bec0305d1
989403855dfa2/t_o_eso_users.jar
2022-01-28 14:37:36.691 WARN manager.MySQLManager: It looks like you are importing from mysql.
2022-01-28 14:37:36.691 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
2022-01-28 14:37:36.691 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
2022-01-28 14:37:36.691 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
2022-01-28 14:37:36.716 INFO mapreduce.ImportJobBase: Beginning import of t_o_eso_users
2022-01-28 14:37:36.717 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.j
ob.tracker.address
2022-01-28 14:37:36.815 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
2022-01-28 14:37:36.833 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job
.maps
Exception in thread "main" java.lang.NoSuchMethodError: org.apache.hadoop.hbase.client.HBaseAdmin.<init>(Lorg/apac
he/hadoop/conf/Configuration;V
    at org.apache.sqoop.mapreduce.HBaseImportJob.jobSetup(HBaseImportJob.java:163)
    at org.apache.sqoop.mapreduce.ImportJobBase.runImport(ImportJobBase.java:268)
    at org.apache.sqoop.manager.SqlManager.importTable(SqlManager.java:692)
    at org.apache.sqoop.manager.MySQLManager.importTable(MySQLManager.java:127)
    at org.apache.sqoop.tool.ImportTool.importTable(ImportTool.java:520)
    at org.apache.sqoop.tool.ImportTool.run(ImportTool.java:628)
    at org.apache.sqoop.Sqoop.run(Sqoop.java:147)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
    at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:183)
    at org.apache.sqoop.Sqoop.runTool(Sqoop.java:234)
    at org.apache.sqoop.Sqoop.runTool(Sqoop.java:243)
    at org.apache.sqoop.Sqoop.main(Sqoop.java:252)

```

The following is an example of running the Sqoop command to extract data. There can be security risks if a command contains the authentication password. You are advised to disable the command recording function (history) before running the command.

```
sqoop import \  
--connect jdbc:mysql://MySQLServer address:Port number/database1 \  
--username admin \  
--password xxx \  
--table table1 \  
--hbase-table table2 \  
--column-family info \  
--hbase-row-key id \  
--hbase-create-table --m 1
```

## Procedure

After the Sqoop client is installed, the JAR packages on which HBase depends are not imported. You need to manually import the JAR packages on which HBase of an earlier version depends.

**Step 1** Check whether the Sqoop and HBase clients are in the same path.

- If yes, go to [Step 2](#).
- If no, delete the original Sqoop and HBase client files, download the complete clients from FusionInsight Manager, and install them in the same path. Then go to [Step 2](#).

**Step 2** Log in to the node where the Sqoop client is installed as user **root**.

**Step 3** Download the following JAR packages of HBase 1.6.0 and upload them to the **lib** directory on the Sqoop client:

- [hbase-client-1.6.0.jar](#)
- [hbase-common-1.6.0.jar](#)
- [hbase-protocol-1.6.0.jar](#)
- [hbase-server-1.6.0.jar](#)

**Step 4** After the packages are uploaded, run the following command to change the permission on the packages to **755**:

```
chmod 755 Package name
```

**Step 5** Run the following command in the client directory to refresh the Sqoop client:

```
source bigdata_env
```

Run the target Sqoop command again.

----End

## 20.3 An Error Is Reported When a Sqoop Task Is Created Using Hue to Import Data from HBase to HDFS

### Issue

An error is reported when a Sqoop operation is performed on Hue to export data from HBase to HDFS.

```
Caused by: java.lang.ClassNotFoundException: org.apache.htrace.Trace
```



This section applies only to MRS 1.9.2 clusters.

```

2022-03-02 15:09:00,264 [main] ERROR org.apache.sqoop.connector.hbase.HBaseExtractor - An exceptional condition has occurred.
org.apache.sqoop.common.SqoopException: HBASE_CONNECTOR_0011:Failed to open table.
    at org.apache.sqoop.connector.hbase.HBaseExtractor.openDB(HBaseExtractor.java:239)
    at org.apache.sqoop.connector.hbase.HBaseExtractor.access$100(HBaseExtractor.java:34)
    at org.apache.sqoop.connector.hbase.HBaseExtractor$1.run(HBaseExtractor.java:86)
    at org.apache.sqoop.connector.hbase.HBaseExtractor$1.run(HBaseExtractor.java:76)
    at org.apache.sqoop.connector.hbase.HBaseExtractor.extract(HBaseExtractor.java:114)
    at org.apache.sqoop.connector.hbase.HBaseExtractor.extract(HBaseExtractor.java:34)
    at org.apache.sqoop.job.mr.SqoopMapper.runInternal(SqoopMapper.java:156)
    at org.apache.sqoop.job.mr.SqoopMapper.run(SqoopMapper.java:79)
    at org.apache.hadoop.mapred.MapTask.runNewMapper(MapTask.java:787)
    at org.apache.hadoop.mapred.MapTask.run(MapTask.java:341)
    at org.apache.hadoop.mapred.YarnChild$2.run(YarnChild.java:188)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:422)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1840)
    at org.apache.hadoop.mapred.YarnChild.main(YarnChild.java:182)
Caused by: java.lang.reflect.InvocationTargetException
    
```

```

Caused by: java.lang.reflect.InvocationTargetException
    at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
    at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:62)
    at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
    at java.lang.reflect.Constructor.newInstance(Constructor.java:423)
    at org.apache.hadoop.hbase.client.ConnectionFactory.createConnection(ConnectionFactory.java:238)
    at org.apache.hadoop.hbase.client.ConnectionManager.createConnection(ConnectionManager.java:454)
    at org.apache.hadoop.hbase.client.ConnectionManager.createConnection(ConnectionManager.java:447)
    at org.apache.hadoop.hbase.client.ConnectionManager.getConnectionInternal(ConnectionManager.java:325)
    at org.apache.hadoop.hbase.client.HTable.<init>(HTable.java:184)
    at org.apache.hadoop.hbase.client.HTable.<init>(HTable.java:150)
    at org.apache.sqoop.connector.hbase.HBaseExtractor.openDB(HBaseExtractor.java:236)
    ... 14 more
Caused by: java.lang.NoClassDefFoundError: org/apache/htrace/Trace
    at org.apache.hadoop.hbase.zookeeper.RecoverableZooKeeper.exists(RecoverableZooKeeper.java:245)
    at org.apache.hadoop.hbase.zookeeper.ZKUtil.checkExists(ZKUtil.java:436)
    at org.apache.hadoop.hbase.zookeeper.ZKClusterId.readClusterIdNode(ZKClusterId.java:65)
    at org.apache.hadoop.hbase.client.ZooKeeperRegistry.getClusterId(ZooKeeperRegistry.java:105)
    at org.apache.hadoop.hbase.client.ConnectionManager$HConnectionImplementation.retrieveClusterId(ConnectionManager.java:944)
    at org.apache.hadoop.hbase.client.ConnectionManager$HConnectionImplementation.<init>(ConnectionManager.java:728)
    ... 25 more
Caused by: java.lang.ClassNotFoundException: org.apache.htrace.Trace
    at java.net.URLClassLoader.findClass(URLClassLoader.java:382)
    at java.lang.ClassLoader.loadClass(ClassLoader.java:419)
    at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:353)
    
```

## Symptom

The Sqoop task is executed successfully, but the CSV file in HDFS is empty.

Name	Description	Creator	Activation	Last Execution	Use Time	Progress	Status	Operate
<input type="checkbox"/>	hbaseToHdfs	hbaseTest->hdfsTest	admin	Enabled	2022/03/02 15:09:04	33s	100%	SUCCEEDED

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r-----	loader	hadoop	0 B	Mar 02 15:09	3	128 MB	hbaseToHdfs-2022-03-02_15.09.00.121.csv

## Cause Analysis

JAR packages conflict or related JAR packages are missing.

## Procedure

**Step 1** Go to the Sqoop installation directory and search for files.



```
[root@node-master1PMPi lib]# cat /etc/hosts
::1      localhost        localhost.localdomain localhost6
127.0.0.1 localhost        localhost.localdomain localh
127.0.0.1 image-pipeline-1004600 image-pipeline-1004600
127.0.0.1 ecs-73f1-191-base ecs-73f1-191-base
1.1.1.1  hadoop.d0edba23_74ce_4527_9e04_22bc21853bb9.com
1.1.1.1  hadoop.hadoop.com
1.1.1.1  hacluster
1.1.1.1  haclusterX
1.1.1.1  haclusterX1
1.1.1.1  haclusterX2
1.1.1.1  haclusterX3
1.1.1.1  haclusterX4
1.1.1.1  ClusterX
1.1.1.1  manager
192.168.9.152 node-master1PMPi.mrs-5s0w.com
192.168.9.200 node-ana-coretmqV.mrs-5s0w.com
[root@node-master1PMPi lib]#
```

7. Run the Sqoop task again. The following error information is displayed.

```
at java.lang.Thread.run(Thread.java:46)
Caused by: com.google.protobuf.ServiceException: java.lang.NoClassDefFoundError: com/yammer/metrics/core/Gauge
at org.apache.hadoop.hbase.ipc.AbstractRpcClient.callBlockingMethod(AbstractRpcClient.java:240)
at org.apache.hadoop.hbase.ipc.AbstractRpcClient$BlockingRpcChannelImplementation.callBlockingMethod(AbstractR
at org.apache.hadoop.hbase.protobuf.generated.ClientProtos$ClientService$BlockingStub.scan(ClientProtos.java:3E
at org.apache.hadoop.hbase.client.ClientSmallReversedScanner$SmallReversedScannerCallable.call(ClientSmallRever
... 9 more
Caused by: java.lang.NoClassDefFoundError: com/yammer/metrics/core/Gauge
at org.apache.hadoop.hbase.ipc.AbstractRpcClient.callBlockingMethod(AbstractRpcClient.java:225)
... 12 more
Caused by: java.lang.ClassNotFoundException: com.yammer.metrics.core.Gauge
at java.net.URLClassLoader.findClass(URLClassLoader.java:362)
at java.lang.ClassLoader.loadClass(ClassLoader.java:419)
at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:352)
at java.lang.ClassLoader.loadClass(ClassLoader.java:352)
... 13 more
2022-03-03 15:45:01,714 [main] INFO org.apache.sqoop.job.mr.SqoopMapper - Extractor has finished
2022-03-03 15:45:01,715 [main] INFO org.apache.sqoop.job.mr.SqoopMapper - Stopping progress service
2022-03-03 15:45:01,727 [main] INFO org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor - SqoopOutputFormatLoadExec
2022-03-03 15:45:01,776 [OutputFormatLoader-consumer] INFO org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor - Lc
2022-03-03 15:45:01,777 [main] INFO org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor - SqoopOutputFormatLoadExec

Log Type: stdout
Log Upload Time: Thu Mar 03 15:45:15 +0800 2022
Log Length: 0

Log Type: syslog
```

**Step 2** Go to the HBase installation directory and search for files.

1. Go to the **lib** directory of HBase and run **grep**.

```
[root@node-master1PMPi lib]# pwd
/opt/Bigdata/MRS_1.9.2/install/FusionInsight-HBase-1.3.1/hbase/lib
[root@node-master1PMPi lib]# grep com.yammer.metrics.core.Gauge *
grep: jline: Is a directory
Binary file metrics-core-2.2.0.jar matches
grep: native: Is a directory
grep: ruby: Is a directory
grep: ruby_luna: Is a directory
[root@node-master1PMPi lib]#
```

2. Copy the JAR package.

```
cp /opt/Bigdata/MRS_1.9.2/install/FusionInsight-HBase-1.3.1/hbase/lib/
metrics-core-2.2.0.jar /opt/Bigdata/MRS_1.9.2/install/FusionInsight-
Hadoop-2.8.3/hadoop/share/hadoop/common/lib/
```

3. Change permissions.

```
chmod 777 metrics-core-2.2.0.jar (the copied JAR package)
```

```
chown omm:ficommon metrics-core-2.2.0.jar (the copied JAR package)
```

4. View the **hosts** file and perform the same operations to copy the JAR package for all other nodes.
5. Run the Sqoop task.

```

2022-03-03 15:50:16,923 INFO [main] org.apache.zookeeper.ZooKeeper: Session: 0xf0000078a340c58 closed
2022-03-03 15:50:16,924 INFO [main-EventThread] org.apache.zookeeper.ClientConn: EventThread shut down for session: 0xf0000078a340c58
2022-03-03 15:50:16,934 INFO [main] org.apache.sqoop.job.mr.SqoopMapper: Extractor has finished
2022-03-03 15:50:16,935 INFO [main] org.apache.sqoop.job.mr.SqoopMapper: Stopping progress service
2022-03-03 15:50:16,942 INFO [main] org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor: SqoopOutputFormatLoadExecutor: SqoopRecordWriter is about to be closed
2022-03-03 15:50:17,397 INFO [OutputFormatLoader-consumer] org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor: Loader has finished
2022-03-03 15:50:17,398 INFO [main] org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor: SqoopOutputFormatLoadExecutor: SqoopRecordWriter is closed
2022-03-03 15:50:17,398 INFO [main] org.apache.hadoop.mapred.Task: Task:attempt_1646292920879_0002_m_000000_0 is done. And is in the process of committing
2022-03-03 15:50:17,435 INFO [main] org.apache.hadoop.mapred.Task: Task:attempt_1646292920879_0002_m_000000_0 done.
2022-03-03 15:50:17,437 INFO [main] org.apache.hadoop.mapred.Task: Final Counters for attempt_1646292920879_0002_m_000000_0: Counters: 26

File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=662083
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=107
  HDFS: Number of bytes written=10
  HDFS: Number of read operations=1
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=1

Map-Reduce Framework
  Map input records=0
  Map output records=1
  Input split bytes=107
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=239
  CPU time spent (ms)=2200
  Physical memory (bytes) snapshot=669523968
  Virtual memory (bytes) snapshot=2691564160
  Total committed heap usage (bytes)=60834648

File Input Format Counters
  Bytes Read=0

File Output Format Counters
  Bytes Written=0
org.apache.sqoop.submission.counter.SqoopCounters
  FILES_WRITTEN=1
  BYTES_READ=1
  BYTES_WRITTEN=1
2022-03-03 15:50:17,538 INFO [main] org.apache.hadoop.metrics2.impl.MetricsSystemImpl: Stopping MapTask metrics system.
2022-03-03 15:50:17,538 INFO [main] org.apache.hadoop.metrics2.impl.MetricsSystemImpl: MapTask metrics system stopped.
2022-03-03 15:50:17,538 INFO [main] org.apache.hadoop.metrics2.impl.MetricsSystemImpl: MapTask metrics system shutdown complete.
    
```

----End

## Conclusion

1. Copy **htrace-core-3.1.0-incubating.jar** in the **lib** directory of Sqoop and **metrics-core-2.2.0.jar** in the **lib** directory of HBase to **/opt/Bigdata/MRS\_1.9.2/install/FusionInsight-Hadoop-2.8.3/hadoop/share/hadoop/common/lib/**.
2. Change the permissions for the JAR packages to **777** and **omm:ficommon**, respectively.
3. Perform the preceding operations on all nodes and run the Sqoop task again.

## 20.4 A Data Format Error Is Reported When Data Is Exported from Hive to MySQL 8.0 Using Sqoop

This section applies only to MRS 3.1.0 clusters.

### Issue

A data format error is reported when Sqoop in an MRS 3.1.0 cluster is used to export data from Hive to MySQL 8.0.





```
fields-terminated-by ',' --input-null-string '\\N' --input-null-non-string '\\N' -  
m 1
```

## 20.5 An Error Is Reported When the sqoop import Command Is Executed to Extract Data from PostgreSQL to Hive

### Background

A user runs the **sqoop import** command to extract data from the open-source PostgreSQL database to MRS HDFS or Hive.

### Issue

The **sqoop** command can be executed to query the PostgreSQL database table, but an error is reported when the **sqoop import** command is executed to import data.

The authentication type 5 is not supported. Check that you have configured the `pg_hba.conf` file to include the client's IP address or subnet.

### Cause Analysis

1. MD5 authentication for connecting to PostgreSQL fails. A whitelist needs to be configured in the `pg_hba.conf` file.
2. When the **sqoop import** command is executed, a MapReduce job is started. The PostgreSQL driver package `gsjdb4-*.jar` exists in the MRS Hadoop installation directory, for example, `/opt/Bigdata/FusionInsight_HD_*/1*_NodeManager/install/hadoop/share/hadoop/common/lib`, which is incompatible with the open-source PostgreSQL service. As a result, an error is reported.

### Procedure

**Step 1** Configure a whitelist in the `pg_hba.conf` file.

**Step 2** Delete the `gsjdb4 jar` packages from all core nodes, and add the PostgreSQL JAR package to `sqoop/lib`.

```
mv /opt/Bigdata/FusionInsight_HD_*/1*_NodeManager/install/hadoop/share/  
hadoop/common/lib/gsjdb4-*/tmp
```

```
1$ mv /opt/Bigdata/FusionInsight_HD_8.1.0.1/1_2_NodeManager/install/hadoop/share/hadoop/common/lib/gsjdb4-V100R003C10SPC125.jar /tmp  
1$ exit
```

----End

## 20.6 Failed to Use Sqoop to Read MySQL Data and Write Parquet Files to OBS

### Issue

An error is reported when Sqoop reads MySQL data and writes the data to OBS in Parquet format. However, the data can be successfully written to OBS if the Parquet format is not specified.

### Symptom

```
2022-02-09 16:36:53.393 ERROR Sqoop: Got exception running Sqoop: org.kitesdk.data.DatasetNotFoundException: Unknown dataset URI pattern: dataset:obs://for
mrs/user/hive/warehouse/dws.db/dws_ks_vip_user_valid_member_id/pts=2022-01-09/part-00000-e6a4d458-f01b-4d8d-906d-3b515815811e.c000
Check that JARs for obs datasets are on the classpath
org.kitesdk.data.DatasetNotFoundException: Unknown dataset URI pattern: dataset:obs://formrs/user/hive/warehouse/dws.db/dws_ks_vip_user_valid_member_id/pts=2022
-01-09/part-00000-e6a4d458-f01b-4d8d-906d-3b515815811e.c000
Check that JARs for obs datasets are on the classpath
at org.kitesdk.data.spi.Registration.lookupDatasetUri(Registration.java:128)
at org.kitesdk.data.Datasets.load(Datasets.java:183)
at org.kitesdk.data.Datasets.load(Datasets.java:149)
at org.kitesdk.data.mapreduce.DatasetKeyInputFormat$ConfigBuilder.readFrom(DatasetKeyInputFormat.java:92)
at org.kitesdk.data.mapreduce.DatasetKeyInputFormat$ConfigBuilder.readFrom(DatasetKeyInputFormat.java:139)
at org.apache.sqoop.mapreduce.JdbcExportJob.configureInputFormat(JdbcExportJob.java:63)
at org.apache.sqoop.mapreduce.ExportJobBase.runExport(ExportJobBase.java:434)
at org.apache.sqoop.manager.SqlManager.exportTable(SqlManager.java:931)
at org.apache.sqoop.tool.ExportTool.exportTable(ExportTool.java:88)
at org.apache.sqoop.tool.ExportTool.run(ExportTool.java:99)
at org.apache.sqoop.Sqoop.run(Sqoop.java:147)
at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:183)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:234)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:243)
at org.apache.sqoop.Sqoop.main(Sqoop.java:252)
2022-02-09 16:36:53.398 WARN metrics.OBSMetricsProvider: Fetch slotid failed.
[root@ecs-gateway mrsclient]# sqoop export --connect jdbc:mysql://10.50.160.241:3306/data_market --username root --password xxx --table dws_ks_vip_user_vali
d_member_test_export --export-dir obs://formrs/user/hive/warehouse/dws.db/dws_ks_vip_user_valid_member_id/pts=2022-01-09/part-00000-e6a4d458-f01b-4d8d-906d-3b515
815811e.c000 --fields-terminated-by '\t' --
```

### Cause Analysis

The Parquet format does not support Hive 3. Data can be written using HCatalog.

### Procedure

Use HCatalog to write data: Specify the Hive database and table in parameters and modify the SQL statement in the script. There can be security risks if a command contains the authentication password. You are advised to disable the command recording function (history) before running the command.

Original script:

```
sqoop import --connect 'jdbc:mysql://10.160.5.65/huawei_pos_online_00?
zeroDateTimeBehavior=convertToNull' --username root --password xxx
--split-by id
--num-mappers 2
--query 'select * from pos_remark where 1=1 and $CONDITIONS'
--target-dir obs://za-test/dev/huawei_pos_online_00/pos_remark
--delete-target-dir
--null-string '\N'
--null-non-string '\N'
--as-parquetfile
```

Modified script:

```
sqoop import --connect 'jdbc:mysql://10.160.5.65/huawei_pos_online_00?
zeroDateTimeBehavior=convertToNull' --username root --password xxx
--split-by id
--num-mappers 2
--query 'select
id,pos_case_id,pos_transaction_id,remark,update_time,update_user,is_deleted,creator,modifier,gmt_created,g
mt_modified,update_user_id,tenant_code from pos_remark where 1=1 and $CONDITIONS'
--hcatalog-database huawei_dev
--hcatalog-table ods_pos_remark
```

## 20.7 An Error Is Reported When Database Data Is Migrated Using Sqoop

### Issue

- The following error is reported when database data is migrated using MRS Sqoop:  
Communications link failure;  
The driver has not received any packets from the server;
- If the migration source is a GaussDB(DWS) database, the following error message is displayed:  
got exception running sqoop .java.lang.RuntimeException, could not load db driver class.

Figure 20-1 Error reported when database data is migrated using Sqoop

```
at org.apache.sqoop.job.mr.SqoopOutputFormatLoadExecutor$ConsumerThread.run(SqoopOutputFormatLoadExecutor.java:506)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: com.mysql.cj.exceptions.CJCommunicationsException: Communications link failure

The last packet sent successfully to the server was 0 milliseconds ago. The driver has not received any packets from the server.
at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:62)
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:423)
at com.mysql.cj.exceptions.ExceptionFactory.createException(ExceptionFactory.java:61)
at com.mysql.cj.exceptions.ExceptionFactory.createException(ExceptionFactory.java:105)
at com.mysql.cj.exceptions.ExceptionFactory.createException(ExceptionFactory.java:151)
at com.mysql.cj.exceptions.ExceptionFactory.createCommunicationsException(ExceptionFactory.java:167)
at com.mysql.cj.protocol.a.NativeSocketConnection.connect(NativeSocketConnection.java:89)
at com.mysql.cj.NativeSession.connect(NativeSession.java:120)
at com.mysql.cj.jdbc.ConnectionImpl.connectOneTryOnly(ConnectionImpl.java:948)
at com.mysql.cj.jdbc.ConnectionImpl.createNewIO(ConnectionImpl.java:818)
at com.mysql.cj.jdbc.ConnectionImpl.<init>(ConnectionImpl.java:448)
at com.mysql.cj.jdbc.ConnectionImpl.getInstance(ConnectionImpl.java:241)
at com.mysql.cj.jdbc.NonRegisteringDriver.connect(NonRegisteringDriver.java:198)
at java.sql.DriverManager.getConnection(DriverManager.java:1664)
at java.sql.DriverManager.getConnection(DriverManager.java:208)
at org.apache.sqoop.connector.jdbc.GenericJdbcExecutor.<init>(GenericJdbcExecutor.java:125)
at org.apache.sqoop.connector.jdbc.Writer.JdbcObjWriter.<init>(JdbcObjWriter.java:23)
at org.apache.sqoop.connector.jdbc.Writer.JdbcCsvWriter.<init>(JdbcCsvWriter.java:10)
at org.apache.sqoop.connector.jdbc.Writer.WriterFactory.getWriter(WriterFactory.java:28)
```

Figure 20-2 Error reported when the migration source is GaussDB(DWS)

```
(via --connection-manager). Sqoop is going to fall back to org.apache.sqoop.manager.GenericJdbcManager. Please specify explicitly which connection manager should be used next time.
2021-12-23 16:22:53.122 INFO manager.SqlManager: Using default fetchSize of 1000
2021-12-23 16:22:53.122 INFO tool.CodeGenTool: Beginning code generation
2021-12-23 16:22:53.143 ERROR sqoop.Sqoop: got exception running sqoop: java.lang.RuntimeException: Could not load db driver class: com. ...., gauss200
jdbc
java.lang.RuntimeException: Could not load db driver class: com. ...., gauss200.jdbc
at org.apache.sqoop.manager.SqlManager.makeConnection(SqlManager.java:875)
at org.apache.sqoop.manager.GenericJdbcManager.getConnection(GenericJdbcManager.java:59)
at org.apache.sqoop.manager.SqlManager.execute(SqlManager.java:763)
at org.apache.sqoop.manager.SqlManager.execute(SqlManager.java:786)
at org.apache.sqoop.manager.SqlManager.getColumnInfoForRawQuery(SqlManager.java:289)
at org.apache.sqoop.manager.SqlManager.getColumnTypesForRawQuery(SqlManager.java:260)
at org.apache.sqoop.manager.SqlManager.getColumnTypesForQuery(SqlManager.java:253)
at org.apache.sqoop.manager.ConnManager.getColumnTypes(ConnManager.java:336)
at org.apache.sqoop.orm.ClassWriter.getColumnTypes(ClassWriter.java:1872)
at org.apache.sqoop.orm.ClassWriter.generate(ClassWriter.java:1671)
at org.apache.sqoop.tool.CodeGenTool.generateRWM(CodeGenTool.java:106)
at org.apache.sqoop.tool.ImportTool.ImportTable(ImportTool.java:501)
at org.apache.sqoop.tool.ImportTool.run(ImportTool.java:628)
at org.apache.sqoop.Sqoop.run(Sqoop.java:147)
at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:133)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:234)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:243)
at org.apache.sqoop.Sqoop.main(Sqoop.java:252)
```

### Cause Analysis

According to the preceding error information, the possible cause is that the JDBC driver is not correctly loaded on the server or the network is faulty. During data migration using Sqoop, MapReduce jobs run on analysis core nodes. Therefore, ensure that the driver class is properly loaded and the network communication between the source database and cluster nodes is normal.

## Procedure

1. Check whether the source database port is enabled for all cluster nodes. If EIPs are used, EIPs must be bound to all cluster nodes, especially the analysis nodes, to ensure normal job running.

### NOTE

Generally, Sqoop instances are distributed on the master node, but jobs are running on the analysis nodes. Therefore, creating a database connection is to communicate with the master node. Jobs are distributed to the analysis nodes through Yarn tasks.

2. Check whether the JAR file of the database driver exists in the installation directory of the Sqoop instances. Ensure that the path of the JAR file is correct, the file permission, owner, and owner group are correct, and the configuration file is correctly configured. Then, restart the Sqoop instances. In this way, the driver can be loaded properly.
  - Driver path for MRS 3.x clusters: **`/opt/Bigdata/FusionInsight_Current/1_xx LoaderServer/install/FusionInsight-Sqoop-1.99.3/server/webapps/loader/WEB-INF/ext-lib/`**
  - Driver path for versions earlier than MRS 3.x: **`/opt/Bigdata/MRS_XXX/install/FusionInsight-Sqoop-1.99.7/FusionInsight-Sqoop-1.99.7/server/jdbc/`**

### NOTE

For versions earlier than 3.x, you need to modify the configuration file.

3. If you submit jobs using the CLI, you are advised to specify the **--driver** parameter.

If the data source is GaussDB(DWS), add **--driver com.xxx.gauss200.jdbc.Driver** to the CLI.

**Figure 20-3** Configuration of the **--driver** parameter when the data source is GaussDB(DWS)

```
[root@bigdata-hadoop-node-master30yot11b]# sqoop import -Dorg.apache.sqoop.splitter.allow_text_splitter=true \  
> --connect jdbc:gaussdb://192.168.100.100:8000/postgres \ --driver com.huawei.gauss200.jdbc.Driver \  
> --username huawei \  
> --password huawei123456 \  
> --query 'SELECT id,created_date,city_name,lat,lon,replace(replace(raw_text,chr(10),''),chr(13),'') as raw_text,service_id,replace(replace(slot_json,chr(10),''),chr(13),'') as slot_json,slot_type,time,wstext,app_verstion,vin,app_platform_entity_id,model_series,ls_test from voicedata.ods_talk_inport_transfer where \${CONDITIONS}' \  
> --delete-target-dir \  
> --target-dir obs://bigdata-sc/databases_dev/ghy_ods_dev.db/ods_talk_import_transfer \  
> --fields-terminated-by '\001' \  
> -m 1
```

For other databases, configure the **--driver** parameter based on site requirements.

# 21 Using Storm

---

## 21.1 Invalid Hyperlink of Events on the Storm Web UI

### Issue

The hyperlink of events on the Storm UI is invalid.

### Symptom

After submitting a topology, a user cannot view topology data processing logs and the events hyperlink is invalid.

### Cause Analysis

The function of viewing topology data processing logs is disabled by default when a topology is submitted in an MRS cluster.

### Procedure

**Step 1** Go to the service page.

- For versions earlier than MRS 3.x: Click the cluster name on the MRS console and choose **Components**.

 **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 3.x or later: Log in to FusionInsight Manager and choose **Cluster**. Click the name of the desired cluster and choose **Services**.

**Step 2** Log in to the Storm web UI.

- For MRS 2.x and earlier versions: Choose **Storm**. On the **Storm WebUI** page, click any UI link to open the Storm web UI.

 NOTE

When accessing the Storm web UI for the first time, you must add the address to the trusted site list.

- For MRS 3.x or later: Choose **Storm > Overview**. On the **Storm WebUI** in the **Basic Information** area, click any UI link to open the Storm web UI.

**Step 3** In the **Topology Summary** area, click the desired topology to view details.

**Step 4** In the **Topology actions** area, click **Kill** to delete the submitted Storm topology.

**Step 5** Submit the Storm topology again and enable the function of viewing topology data processing logs. Add the **topology.eventlogger.executors** parameter and set it to a positive integer when submitting the Storm topology. Example:

```
storm jar Path of the topology package Class name of the topology Main method  
Topology name -c topology.eventlogger.executors=X
```

**Step 6** In the **Topology Summary** area on the Storm UI, click the desired topology to view details.

**Step 7** In the **Topology actions** area, click **Debug**, specify the data sampling percentage, and click **OK**.

**Step 8** Click the **Spouts** or **Bolts** task name of the topology. In **Component summary**, click **events** to view data processing logs.

 NOTE

To enable the function of viewing topology data processing logs of the specified **Spouts** or **Bolts** task, click the **Spouts** or **Bolts** task name of the topology, click **Debug** in the **Topology actions** area, and enter the data sampling percentage.

----End

## 21.2 Failed to Submit the Storm Topology

### Symptom

An MRS streaming cluster is installed, and ZooKeeper, Storm, and Kafka are installed in the cluster.

A topology fails to be submitted by running commands on the client.

### Possible Causes

- The Storm service is abnormal.
- The client user is not authenticated or the authentication has expired.
- The **storm.yaml** file in the submitted topology conflicts with that on the server.

### Cause Analysis

A user fails to submit the topology. The possible cause is that the client or Storm is faulty.

1. Check the Storm status.  
For versions earlier than MRS 3.x: Log in to MRS Manager. On the MRS Manager page, choose **Services > Storm** to check the status of Storm. The status is **Good**, and the monitoring metrics are correctly displayed.  
For MRS 3.x or later: Log in to FusionInsight Manager. Choose **Cluster > Services > Storm** to check the status of Storm. It is found that the status is **Good** and the monitoring metrics are correctly displayed.
2. Check the submission logs of the client. The logs contain "KeeperExceptionSessionExpireException".

```
org.apache.zookeeper.KeeperException$SessionExpiredException: KeeperErrorCode = Session expired
at org.apache.zookeeper.KeeperException.create(KeeperException.java:131) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.curator.frameworkimps.CuratorFrameworkImpl.checkBackgroundRetry(CuratorFrameworkImpl.java:710) [curator-framework-2.5.0.jar:na]
at org.apache.curator.frameworkimps.CuratorFrameworkImpl.processBackgroundOperation(CuratorFrameworkImpl.java:510) [curator-framework-2.5.0.jar:na]
at org.apache.curator.frameworkimps.BackgroundSyncImpl1.processResult(BackgroundSyncImpl1.java:150) [curator-framework-2.5.0.jar:na]
at org.apache.zookeeper.ClientCnxn$EventThread.processEvents(ClientCnxn.java:484) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.zookeeper.ClientCnxn$EventThread.queuePacket(ClientCnxn.java:498) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.zookeeper.ClientCnxn.finishPacket(ClientCnxn.java:731) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.zookeeper.ClientCnxn.sendPacket(ClientCnxn.java:745) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.zookeeper.ClientCnxn.access$2700(ClientCnxn.java:97) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.zookeeper.ClientCnxn$SendThread.cleanup(ClientCnxn.java:1391) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
at org.apache.zookeeper.ClientCnxn$SendThread.run(ClientCnxn.java:1314) [zookeeper-3.5.0.jar:3.5.0-V1008002C00B109]
2016-05-31 09:13:24 | INFO | [main] | Session: 0x100273947405ab4b closed | org.apache.zookeeper.ZooKeeper (ZooKeeper.java:948)
Exception in thread "main" java.lang.RuntimeException: Exception while initializing NimbusLeaderElections
at backtype.storm.nimbus.NimbusLeaderElections.init(NimbusLeaderElections.java:84)
at backtype.storm.utils.NimbusClient.getConfiguredClient(NimbusClient.java:39)
at backtype.storm.StormSubmitter.submitTopology(StormSubmitter.java:159)
at backtype.storm.StormSubmitter.submitTopologyWithProgressBar(StormSubmitter.java:254)
at backtype.storm.StormSubmitter.submitTopologyWithProgressBar(StormSubmitter.java:236)
at storm.starter.WordCountTopology.main(WordCountTopology.java:94)
Caused by: org.apache.zookeeper.KeeperException$ConnectionLossException: KeeperErrorCode = ConnectionLoss for /storm/nimbus-leader
at org.apache.zookeeper.KeeperException.create(KeeperException.java:99)
at org.apache.zookeeper.KeeperException.create(KeeperException.java:81)
at org.apache.zookeeper.ZooKeeper.exists(ZooKeeper.java:1501)
at org.apache.curator.frameworkimps.ExistsBuilderImpl162.call(ExistsBuilderImpl162.java:172)
at org.apache.curator.frameworkimps.ExistsBuilderImpl162.call(ExistsBuilderImpl162.java:161)
at org.apache.curator.RetryLoop.callWithRetry(RetryLoop.java:107)
at org.apache.curator.frameworkimps.ExistsBuilderImpl162.pathInForeground(ExistsBuilderImpl162.java:157)
at org.apache.curator.frameworkimps.ExistsBuilderImpl162.forPath(ExistsBuilderImpl162.java:148)
at org.apache.curator.frameworkimps.ExistsBuilderImpl162.forPath(ExistsBuilderImpl162.java:146)
at backtype.storm.nimbus.NimbusLeaderElections.init(NimbusLeaderElections.java:64)
... 5 more
```

The preceding error occurs because security authentication is not performed before the topology is submitted or the TGT expires after authentication.

For details about the solution, see [Step 1](#).

3. Check the client submission log. It is found that the "ExceptionInInitializerError" exception information is printed, and the message "Found multiple storm.yaml resources" is displayed. The following is an example:

```
Exception in thread "main" java.lang.ExceptionInInitializerError
at backtype.storm.topology.TopologyBuilder.createTopology(TopologyBuilder.java:106)
at com.huawei.streaming.storm.example.wordcount.WordCountTopology.cmdSubmit(WordCountTopology.java:117)
at com.huawei.streaming.storm.example.wordcount.WordCountTopology.submitTopology(WordCountTopology.java:80)
at com.huawei.streaming.storm.example.wordcount.WordCountTopology.main(WordCountTopology.java:71)
Caused by: java.lang.RuntimeException: Found multiple storm.yaml resources. You're probably bundling the Storm jars with your topology jar.
at backtype.storm.utils.Utils.findAndReadConfigFile(Utils.java:151)
at backtype.storm.utils.Utils.readStormConfig(Utils.java:206)
at backtype.storm.utils.Utils.<clinit>(Utils.java:70)
... 4 more
```

This error occurs because the **storm.yaml** file in the service JAR package conflicts with that on the server.

For details about the solution, see [Step 2](#).

4. If the fault is not caused by the preceding reasons, see [Failed to Submit the Storm Topology and Message "Failed to check principle for keytab" Is Displayed](#).

## Solution

**Step 1** An authentication error occurs.

1. Log in to the node where the client resides and switch to the client directory.
2. Run the following command to submit the task again: (Replace the service JAR package and topology based on the site requirements.)

```
source bigdata_env
kinit Username
```

**storm jar storm-starter-topologies-0.10.0.jar  
storm.starter.WordCountTopology test**

**Step 2** The topology package is abnormal.

Check the service JAR package, delete the **storm.yaml** file from the service JAR package, and submit the task again.

----End

## 21.3 Failed to Submit the Storm Topology and Message "Failed to check principle for keytab" Is Displayed

### Symptom

An MRS streaming cluster in security mode is installed, and ZooKeeper, Storm, and Kafka are installed in the cluster.

When a topology is defined to access components such as HDFS and HBase and the topology fails to be submitted using client commands.

### Possible Causes

- The submitted topology does not contain the keytab file of the user.
- The keytab file contained in the submitted topology is inconsistent with the user who submits the topology.
- The **user.keytab** file exists in the **/tmp** directory on the client, and the owner is not the running user.

### Cause Analysis

1. Check the logs. Error information "Can not found user.keytab in storm.jar" is found. Details are as follows:

```
[main] INFO b.s.StormSubmitter - Get principle for stream@HADOOP.COM success  
[main] ERROR b.s.StormSubmitter - Can not found user.keytab in storm.jar.  
Exception in thread "main" java.lang.RuntimeException: Failed to check principle for keytab  
at backtype.storm.StormSubmitter.submitTopologyAs(StormSubmitter.java:219)
```

Check the JAR file of the submitted topology. It is found that the keytab file is not contained.

2. Check the logs. Error information "The submit user is invalid,the principle is" is found. Details are as follows:

```
[main] INFO b.s.StormSubmitter - Get principle for stream@HADOOP.COM success  
[main] WARN b.s.s.a.k.ClientCallbackHandler - Could not login: the client is being asked for a  
password, but the client code does not currently support obtaining a password from the user. Make  
sure that the client is configured to use a ticket cache (using the JAAS configuration setting  
'useTicketCache=true') and restart the client. If you still get this message after that, the TGT in the  
ticket cache has expired and must be manually refreshed. To do so, first determine if you are using a  
password or a keytab. If the former, run kinit in a Unix shell in the environment of the user who is  
running this client using the command 'kinit <princ>' (where <princ> is the name of the client's  
Kerberos principal). If the latter, do 'kinit -k -t <keytab> <princ>' (where <princ> is the name of the  
Kerberos principal, and <keytab> is the location of the keytab file). After manually refreshing your  
cache, restart this client. If you continue to see this message after manually refreshing your cache,  
ensure that your KDC host's clock is in sync with this host's clock.  
[main] ERROR b.s.StormSubmitter - The submit user is invalid,the principle is : stream@HADOOP.COM  
Exception in thread "main" java.lang.RuntimeException: Failed to check principle for keytab  
at backtype.storm.StormSubmitter.submitTopologyAs(StormSubmitter.java:219)
```



The authenticated user used to submit the topology is **stream**. However, the system displays a message indicating that the submit user is invalid during topology submission, indicating that the internal verification fails.

3. Check the JAR file of the submitted topology. It is found that the keytab file is contained.

The **principal** is set to **zmk\_kafka** in the **user.keytab** file.

```
[root@8-5-148-6 client]# klist -kt user.keytab
Keytab name: FILE:user.keytab
KVNO Timestamp Principal
-----
1 12/19/16 16:28:17 zmk_kafka@HADOOP.COM
1 12/19/16 16:28:17 zmk_kafka@HADOOP.COM
```

It is found that the authenticated user does not match the principal in the **user.keytab** file.

4. Check the logs and find the error information "Delete the tmp keytab file failed, the keytab file is:/tmp/user.keytab". The detailed information is as follows:

```
[main] WARN b.s.StormSubmitter - Delete the tmp keytab file failed, the keytab file is : /tmp/
user.keytab
[main] ERROR b.s.StormSubmitter - The submit user is invalid,the principle is : hbase1@HADOOP.COM
Exception in thread "main" java.lang.RuntimeException: Failed to check principle for keytab
at backtype.storm.StormSubmitter.submitTopologyAs(StormSubmitter.java:213)
```

Check the **/tmp** directory. It is found that the **user.keytab** file exists and the file owner is not the running user.

## Solution

- Ensure that the **user.keytab** file is carried when the topology is submitted.
- Ensure that the user for submitting the topology is the same as that of the **user.keytab** file.
- Delete the **user.keytab** file from the **/tmp** directory.

## 21.4 Worker Logs Are Empty After the Storm Topology Is Submitted

### Symptom

After a topology is remotely submitted in Eclipse, the detailed information about the topology cannot be viewed on the Storm web UI, and the Worker node where Bolt and Spout of each topology are located keeps changing. The Worker log is empty.

### Possible Causes

The Worker process fails to be started, triggering Nimbus to re-allocate tasks and start the Worker process on other Supervisors. The Worker process continues to restart. As a result, the Worker node keeps changing, and the Worker log is empty. The possible causes of the Worker process startup failure are as follows:

- The submitted JAR package contains the **storm.yaml** file.  
Storm specifies that each classpath can contain only one **storm.yaml** file. If there is more than one **storm.yaml** file, an exception occurs. Use the Storm

client to submit the topology. The classpath configuration of the client is different from the classpath configuration of Eclipse. The client automatically loads the JAR package of the user to classpath. As a result, two **storm.yaml** files exist in classpath.

- The initialization of the Worker process takes a long time, which exceeds the Worker startup timeout period set in the Storm cluster. As a result, the Worker process is killed and reallocated.

## Troubleshooting Process

1. Use the Storm client to submit the topology and check whether the **storm.yaml** file is duplicate.
2. Repack the JAR file and submit the topology again.
3. Modify the Worker startup timeout parameter in the Storm cluster.

## Procedure

- Step 1** If the Worker log is empty after the topology is remotely submitted using Eclipse, use the Storm client to submit the JAR package corresponding to the topology and view the prompt message.

For example, if the JAR package contains two **storm.yaml** files in different paths, the following information is displayed:

```
Exception in thread "main" java.lang.ExceptionInInitializerError
  at com.XXXexample.WordCountTopology.createConf(WordCountTopology.java:132)
  Caused by: java.lang.RuntimeException: Found multiple storm.yaml resources. You're probably bundling the
  Storm jars with your topology jar. [jar:file:/XXX/streaming-0.9.2/bin/stormDemo.jar!/storm.yaml, file:/XXX/
  Streaming/streaming-0.9.2/conf/storm.yaml]
  at backtype.storm.utils.Utils.findAndReadConfigFile(Utils.java:151)
```

- Step 2** Compress the JAR package again. Ensure that the package does not contain the **storm.yaml** file and JAR packages related to **log4j** and **slf4j-log4j**.
- Step 3** Use IntelliJ IDEA to remotely submit the new JAR package.
- Step 4** Check whether the topology details and Worker logs can be viewed on the web UI.
- Step 5** On MRS Manager, modify the Worker startup timeout parameter of the Storm cluster (for details about the parameter description, see [Related Information](#)). Save the modification, and restart the Storm service.
- MRS Manager: Log in to MRS Manager and choose **Services > Storm > Configuration**.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Storm > Configurations**.
- Step 6** Submit the JAR package to be run again.

----End

## Related Information

1. The **nimbus.task.launch.secs** and **supervisor.worker.start.timeout.secs** parameters indicate the topology startup timeout tolerance of the Nimbus and supervisor, respectively. Generally, the value of **nimbus.task.launch.secs** must be greater than or equal to that of

**supervisor.worker.start.timeout.secs.** It is recommended that the value of **nimbus.task.launch.secs** be slightly greater or equal to that of **supervisor.worker.start.timeout.secs**. Otherwise, the task reallocation efficiency will be affected.

- **nimbus.task.launch.secs:** If the Nimbus does not receive the heartbeat message sent by the topology task within the period specified by this parameter, the Nimbus re-allocates the topology to another supervisor and updates the task information in ZooKeeper. The supervisor reads the task information in ZooKeeper and compares it with the topology started. If the topology does not belong to the supervisor, the supervisor deletes the metadata of the topology, that is, the `/srv/Bigdata/streaming_data/stormdir/supervisor/stormdist/{worker-id}` directory.
- **supervisor.worker.start.timeout.secs:** After the supervisor starts a worker, if no heartbeat message is received from the worker within the period specified by this parameter, the supervisor stops the worker and waits for worker rescheduling. Generally, the value of this parameter is increased when the service startup takes a long time to ensure that the worker can be started successfully.

If the value of **supervisor.worker.start.timeout.secs** is greater than that of **nimbus.task.launch.secs**, the worker is still started before the tolerance time of supervisor ends. However, the Nimbus considers that the service startup times out and allocates the service to another host. The background thread of the supervisor finds that the tasks are inconsistent and deletes the metadata of the topology. As a result, when the worker attempts to read **stormconf.ser** during startup, the file does not exist, and "FileNotFoundException" is thrown.

2. The **nimbus.task.timeout.secs** and **supervisor.worker.timeout.secs** parameters indicate the timeout tolerance time for the Nimbus and supervisor to report heartbeat messages during topology running. Generally, the value of **nimbus.task.timeout.secs** must be slightly greater than or equal to that of **supervisor.worker.timeout.secs**.

## 21.5 Worker Runs Abnormally After the Storm Topology Is Submitted and Error "Failed to bind to XXX" Is Displayed

### Symptom

After the service topology is submitted, the worker cannot be started normally. Check the worker log. The log records "Failed to bind to: *Host\_ip:Port*".



```
[root@dgg-dggcbgfi056-stm supervisor]#  
[root@dgg-dggcbgfi056-stm supervisor]#  
[root@dgg-dggcbgfi056-stm supervisor]# cat /proc/sys/net/ipv4/ip_local_port_range  
10240 65000  
[root@dgg-dggcbgfi056-stm supervisor]# █
```

5. It is found that the random port range is too large and conflicts with the service port range of MRS.

#### NOTE

The MRS service port number ranges from 20000 to 30000.

## Procedure

- Step 1** Modify the random port range.

```
vi /proc/sys/net/ipv4/ip_local_port_range
```

```
32768 61000
```

- Step 2** Stop the service process that occupies the service port to release the port. (Stop the service topology.)

----End

## 21.6 "well-known file is not secure" Is Displayed When the jstack Command Is Used to Check the Process Stack

### Symptom

When a user runs the **jstack** command to view the process stack information, the following error message is displayed:

```
omm@hadoop02:~> jstack 62517  
62517: well-known file is not secure
```

### Cause Analysis

1. The user running the **jstack** command is inconsistent with the user submitting the process for viewing the pid information.
2. Storm uses the feature of differentiating users for implementing tasks. When the worker process is started, the process UID and GID are changed to the user submitting the task and ficommon. This way, logviewer can access logs of the worker process and only log file permission 640 is open. After the user is changed, the **jstack** and **jmap** commands fail to be executed for the worker process, because the default GID of the user is not ficommon. You need to run the **ldap** command to change the user GID to 9998 (ficommon).

### Solution

You can use either of the following two methods to resolve the problem:

Method 1: View the process stack on the native Storm page.

**Step 1** Log in to the native Storm page.

MRS Manager:

1. Access MRS Manager.
2. Choose **Services > Storm**. In **Storm WebUI** of **Storm Summary**, click any UI link to access the Storm WebUI.

FusionInsight Manager:

1. Log in to FusionInsight Manager.
2. On Manager, choose **Cluster > Service > Storm**. On the **Storm WebUI** page of **Overview**, click any UI link to open the Storm WebUI.

**Step 2** Select the topology to be viewed.

Topology Summary						
Name	Owner	Status	Uptime	Num workers	Num executors	Num tasks
wc	stormuser	ACTIVE	4s	0	0	0

**Step 3** Select the spout or bolt to be viewed.

Spouts (All time)							
Id	Executors	Tasks	Emitted	Transferred	Complete latency (ms)	Acked	Failed
spout	5	5	1500	1500	0.000	0	0

Showing 1 to 1 of 1 entries

Bolts (All time)								
Id	Executors	Tasks	Emitted	Transferred	Capacity (last 10m)	Execute latency (ms)	Executed	Process latency (ms)
count	12	12	13500	0	0.025	0.480	12500	0.160
split	8	8	12500	12500	0.000	0.000	2500	3.000

**Step 4** Select the log file of the node to be viewed, and then click **JStack** or **Heap**. **JStack** corresponds to the stack information, and **Heap** corresponds to the heap information.

Profiling and Debugging							
Use the following controls to profile and debug the components on this page.							
Status / Timeout (Minutes)	Actions						
<input type="text" value="10"/>	<input type="button" value="JStack"/> <input type="button" value="Restart Worker"/> <input type="button" value="Heap"/>						
Executors (All time)							
Id	Uptime	Host	Port	Actions	Emitted	Transferred	Complete latency (ms)
[24-24]	1m 40s	hadoop03	29300	<input checked="" type="checkbox"/> files	1000	1000	0.000
[25-25]	1m 41s	hadoop01	29300	<input type="checkbox"/> files	1000	1000	0.000
[26-26]	1m 41s	hadoop02	29300	<input type="checkbox"/> files	1000	1000	0.000
[27-27]	1m 40s	hadoop03	29300	<input checked="" type="checkbox"/> files	1000	1000	0.000
[28-28]	1m 41s	hadoop01	29300	<input type="checkbox"/> files	1000	1000	0.000

----End

Method 2: View the process stack by modifying user-defined parameters.

**Step 1** Access the Storm parameter configuration page.

MRS Manager: Log in to MRS Manager, choose **Services > Storm > Service Configuration**, and select **All** from the **Type** drop-down list.

Operation on FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Yarn > Configurations > All Configurations**.

**Step 2** In the navigation tree on the left, choose **supervisor > Customize** and add the variable **supervisor.run.worker.as.user=false**.

**Step 3** Click **Save Configuration** and select **Restart the affected services or instances**. Click **OK** to restart the services.

**Step 4** Submit the topology again.

**Step 5** Switch to the **omm** user on the background node and run the **jps** command to view the PID of the worker process.

```
omm@hadoop02:~> jps | grep worker
22485 worker
111402 worker
```

**Step 6** Run the **jstack pid** command to view the jstack information.

```
omm@hadoop02:~> jstack 22485
2018-05-26 08:46:24
Full thread dump Java HotSpot(TM) 64-Bit Server VM (25.144-b01 mixed mode):

"Attach Listener" #82 daemon prio=9 os_prio=0 tid=0x000000001c95000 nid=0xb840 waiting on condition [0x0000000000000000]
 java.lang.Thread.State: RUNNABLE

"pool-14-thread-1" #81 daemon prio=5 os_prio=0 tid=0x000007f7ebc931000 nid=0x6113 waiting on condition [0x000007f7eb5ddf000]
 java.lang.Thread.State: TIMED_WAITING (parking)
   at sun.misc.Unsafe.park(Native Method)
   - parking to wait for <0x00000000dfe820a0> (a java.util.concurrent.locks.AbstractQueuedSynchronizer$ConditionObject)
   at java.util.concurrent.locks.LockSupport.parkNanos(LockSupport.java:215)
   at java.util.concurrent.ScheduledThreadPoolExecutor$DelayedWorkQueue.take(ScheduledThreadPoolExecutor.java:1093)
   at java.util.concurrent.ScheduledThreadPoolExecutor$DelayedWorkQueue.take(ScheduledThreadPoolExecutor.java:809)
   at java.util.concurrent.ThreadPoolExecutor.getTask(ThreadPoolExecutor.java:1074)
   at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1134)
   at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
   at java.lang.Thread.run(Thread.java:748)
```

----End

## 21.7 Data Cannot Be Written to Bolts When the Storm-JDBC Plug-in Is Used to Develop Oracle Databases

### Symptom

When the Storm-JDBC plug-in is used to develop Oracle write Bolts, the Oracle database can be connected, but data cannot be written to the Oracle database.

Bolts (All time)

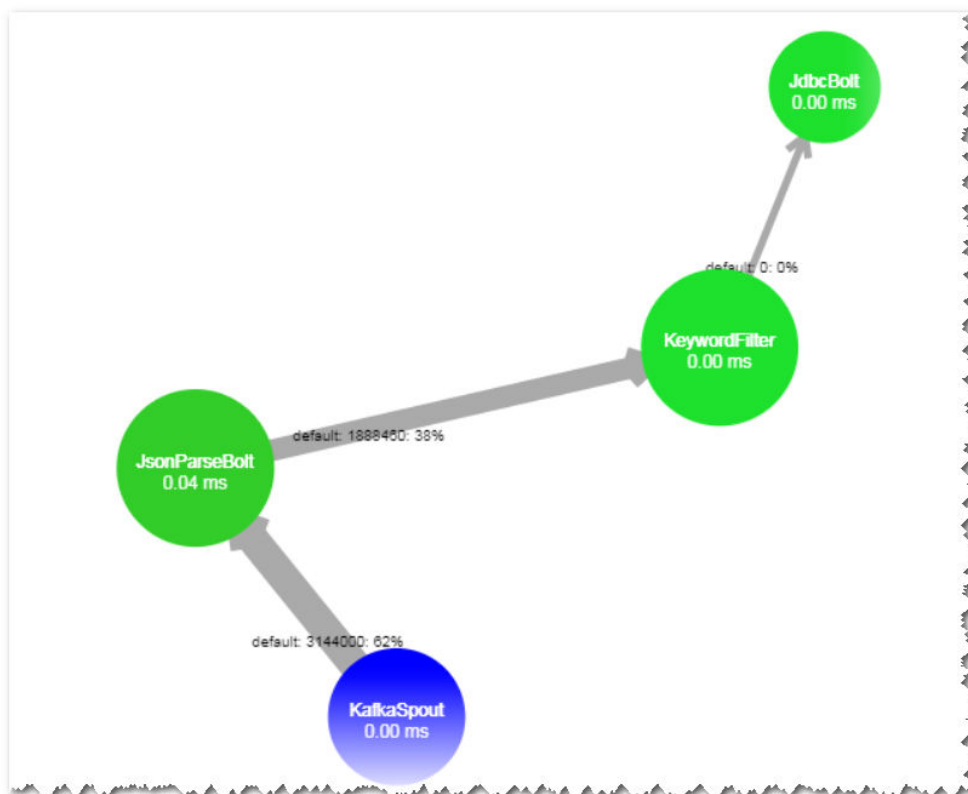
Id	Executors	Tasks	Emitted	Transferred	Capacity (last 10m)	Execute latency (ms)	Executed	Process latency (ms)	Acked	Failed	Error Host	Error Port	Last error
JdbcBolt	2	2	0	0	0.000	0.000	0	0.000	0	0			
JsonParseBolt	5	5	3698140	3698140	0.009	0.048	3700260	0.044	3700200	0			
KeywordFilter	5	5	0	0	0.000	0.001	3592380	0.000	0	0			

## Possible Causes

- The topology definition is incorrect.
- The definition of the database table result is incorrect.

## Cause Analysis

1. On the Storm web UI, check the DAG of the topology. The DAG is consistent with the topology definition.



2. The definition of the KeyWordFilter Bolt is consistent with the **expParser** field.

```

@Override
public void declareOutputFields(OutputFieldsDeclarer declarer)
{
    declarer.declare(new Fields("timestamp", "keyword", "hostname", "message", "kafka_topic" ));
}
  
```

```

if( flag )
{
    String keyword = expParser.getKeyword();
    System.out.println( message );
    collector.emit(new Values( timestamp, keyword , hostname , message, kafka_topic ));
}
  
```

3. View the table definition in the Oracle database. The field name is in uppercase, which is inconsistent with flow definition field name.

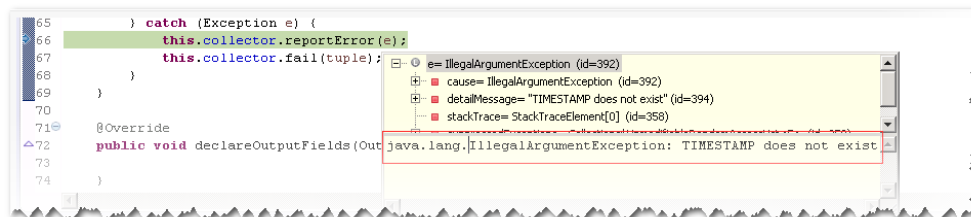
```

select * from KeywordFilterAlarm
  
```

TIMESTAMP	KEYWORD	HOSTNAME	MESSAGE	KAFKA_TOPIC
2018-09-19 10:20:00	contains("china")	jlwasi2-n...	19-Sep-2018 11:20:08 C48 cuenes info client10.200.171.57#6392	dns_query



4. When the execute method is debugged independently, it is found that the thrown field does not exist.



## Procedure

The field name of the stream definition is changed to uppercase letters, which is the same as that defined in the database table.

## 21.8 Internal Server Error Is Displayed When the User Queries Information on the Storm UI

### Symptom

An MRS cluster is installed, and ZooKeeper and Storm are installed in the cluster.

The following information is displayed when a user views information from the Storm status page on MRS Manager:

```
Internal Server Error
org.apache.thrift7.transport.TTransportException: Frame size (306030) larger than max length (1048576)!
```

### Possible Causes

- Nimbus of Storm is abnormal.
- Storm cluster information exceeds the default Thrift transmission size.

### Cause Analysis

1. Check the Storm service status and monitoring metrics:
  - MRS Manager: Log in to MRS Manager and choose **Services > Storm**. Check the Storm status. The status is **Good**, and the monitoring metrics are correctly displayed.
  - FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Storm**. Check the Storm status. The status is **Good**, and the monitoring metrics are correctly displayed.
2. Click the **Instance** tab and check the status of the Nimbus instance. The status is normal.
3. Check the Thrift configuration of the Storm cluster. It is found that **nimbus.thrift.max\_buffer\_size** is set to **1048576** (1 MB).
4. The preceding configuration is the same as that in the exception information, indicating that the buffer size of Thrift is less than that required by the cluster information.

## Procedure

Adjust the Thrift buffer size of the Storm cluster.

**Step 1** Access the Storm parameter configuration page.

- MRS Manager: Log in to MRS Manager, choose **Services > Storm > Service Configuration**, and select **All** from the **Type** drop-down list.
- Operation on FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Yarn > Configurations > All Configurations**.

**Step 2** Change the value of **nimbus.thrift.max\_buffer\_size** to **10485760** (10 MB).

**Step 3** Click **Save Configuration**, select **Restart the affected services or instances**, and click **OK**.

----End

# 22 Using Ranger

---

## 22.1 After Ranger Authentication Is Enabled for Hive, Unauthorized Tables and Databases Can Be Viewed on the Hue Page


### Symptom

In a normal cluster with Kerberos authentication disabled, after Ranger authentication is enabled for Hive, cluster users can view unauthorized data tables and databases on the Hue page.

### Cause Analysis

After Ranger authentication is enabled for Hive, the default Hive policies contain two public group policies about databases. All users belong to the public group. By default, the public group is granted the permission to create tables in the default database and create other databases. Therefore, all users have the **show databases** and **show tables** permissions by default. If some users do not need to have these two permissions, you can delete the default public group policies on the Ranger web UI and grant the required user permissions.

### Procedure

- Step 1** Log in to the Ranger web UI.
- Step 2** In the **Service Manager** area, click the Hive component name to access the Hive security access policy page.
- Step 3** Click  in the rows containing the **all - database** and **default database tables columns** policies.
- Step 4** Delete the public group policies.

**Figure 22-1** all - database policy

Allow Conditions:

Select Role	Select Group	Select User	Policy Conditions	Permissions	Delegate Admin	
Select Roles	Select Groups	in Hive	Add Conditions +	select update create drop Alter insert lock All Read Write RepAdmin Service Admin Temporary UDF Admin Refresh	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Select Roles	in public	Select Users	Add Conditions +	Create	<input type="checkbox"/>	<input type="checkbox"/>
Select Roles	Select Groups	in OWNER	Add Conditions +	All	<input checked="" type="checkbox"/>	<input type="checkbox"/>

**Figure 22-2** default database tables columns policy

Allow Conditions:

Select Role	Select Group	Select User	Policy Conditions	Permissions	Delegate Admin	
Select Roles	in public	Select Users	Add Conditions +	Create	<input type="checkbox"/>	<input type="checkbox"/>
Select Roles	Select Groups	in Hive in OWNER	Add Conditions +	All	<input checked="" type="checkbox"/>	<input type="checkbox"/>

**Step 5** On the Hive security access policy page, click **Add New Policy** to add resource access policies for related users or user groups. For details, see [Configuring Component Permission Policies](#).

----End

# 23 Using Yarn

## 23.1 A Large Number of Jobs Occupying Resources After Yarn Is Started in a Cluster

### Symptom

In an MRS 2.x cluster or earlier, a large number of jobs are generated after Yarn is started, occupying computing resources of the cluster.

The screenshot shows the Yarn cluster metrics page. The 'Cluster Metrics' section displays various statistics, including 'Apps Submitted' (1347), 'Apps Pending' (0), 'Apps Running' (3397), and 'Apps Completed' (0). The 'Cluster Nodes Metrics' section shows 'Nodes' (2), 'Decommissioning Nodes' (0), 'Decommissioned Nodes' (0), 'Lost Nodes' (0), 'Unhealthy Nodes' (0), 'Reserved Nodes' (0), and 'Shutdown Nodes' (0). The 'Scheduler Metrics' section shows 'Scheduler Type' (Capacity Scheduler), 'Scheduling Resource Type' (MEMORY, CPU), 'Minimum Allocation' (memory 512, #Cores 2), 'Maximum Allocation' (memory 51200, #Cores 8), and 'Maximum Cluster Application Priority' (0). The 'Jobs' table shows a large number of failed jobs, with columns for ID, User, Name, Application, Queue, Application Priority, Start Time, Finish Time, State, Final Status, Running Containers, Allocated CPU, Allocated Memory, % of Queue, % of Cluster, Progress, Tracking, and Kill.

### Cause Analysis

If the source IP address of the Any protocol in the inbound direction of the cluster's security group is set to 0.0.0.0/0, the cluster may be attacked from external networks.

IPv4	Any	Any	0.0.0.0/0
IPv4	Any	Any	0.0.0.0/0
IPv4	Any	Any	0.0.0.0/0

## Procedure

- Step 1** Log in to the MRS console. On the **Active Clusters** page, click the cluster name to access the cluster details page.
- Step 2** Click **Access Manager** next to **MRS Manager**. The **Access MRS Manager** page is displayed.
- Step 3** Click **Manage Security Group Rule** to check the security group rule configuration.

### Access MRS Manager

To access MRS Manager, you need to bind an EIP and add security group rules. [Learn more](#)

EIP ?	100.85.217.88	Manage EIP
Security Group	mrs_mrs_ntbg_zha...	
Add Security Group Rule ?	100.79.4.245/32	Manage Security Group Rule

I confirm that 100.79.4.245/32 is a trusted public IP address and MRS Manager can be accessed using this IP address.

OK

Cancel

- Step 4** Check whether the source address of the **Any** protocol in the inbound direction is **0.0.0.0/0**.

If it is, change the remote end of the **Any** protocol in the inbound direction to a specified IP address. If it is not, no change is required.

- Step 5** After the value is changed successfully, restart the cluster.

----End

## Summary and Suggestions

Disable the **Any** protocol in the inbound direction, or set the remote end of the **Any** protocol in the inbound direction to a specified IP address.

## 23.2 Error "GC overhead" Is Reported When Tasks Are Submitted Using the `hadoop jar` Command on the Client

### Symptom

When a user submits a task on the client, the client returns a memory overflow error.

```
main path:hdfs://hacluster/user
17/09/18 08:29:57 INFO hdfs.client: Created HDFS_DELEGATION_TOKEN token 22890097 for wangyou on ha-hdfs:hacluster
17/09/18 08:29:57 INFO security.TokenCache: Got dt for hdfs://hacluster: Kind: HDFS_DELEGATION_TOKEN, Service: ha-hdfs:hacluster, Ident: (HDFS_DELEGATION_TOKEN token 2
2890097) For
17/09/18 08:29:57 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with
ToolRunner to remedy this.
17/09/18 08:32:42 INFO retry.RetryInvocationHandler: Exception while invoking getList of class ClientNameNodeProtocolTranslatorPB over f11-cn-003/10.113.246.10:2500
0. Trying to fail over immediately.
java.io.IOException: com.google.protobuf.ServiceException: java.lang.OutOfMemoryError ([GC overhead limit exceeded]
    at org.apache.hadoop.ipc.ProtobufHelper$GetRemoteException(ProtobufHelper.java:247)
    at org.apache.hadoop.hdfs.protocolPB.ClientNameNodeProtocolTranslatorPB.getListOfClientNameNodeProtocolTranslatorPB(java:1578)
    at sun.reflect.GeneratedMethodAccessor2.invoke(Unknown Source)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:497)
    at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:191)
    at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:102)
    at com.sun.proxy.$Proxy10.getListOf(Unknown Source)
    at org.apache.hadoop.hdfs.DFSClient$1$1.getPath(DFSClient.java:1737)
    at org.apache.hadoop.hdfs.DistributedFileSystem$DistributedFileListingIterator.hasNext(DistributedFileSystem.java:1024)
    at org.apache.hadoop.hdfs.DistributedFileSystem$DistributedFileListingIterator.hasNext(DistributedFileSystem.java:999)
    at org.apache.hadoop.mapreduce.lib.input.FileInputFormat$1$1.setInputFormat(DelegatingInputFormat.java:304)
    at org.apache.hadoop.mapreduce.lib.input.FileInputFormat$1$1.setInputFormat(DelegatingInputFormat.java:285)
    at org.apache.hadoop.mapreduce.lib.input.CombineFileInputFormat$1$1.setInputFormat(CombineFileInputFormat.java:217)
    at org.apache.hadoop.mapreduce.lib.input.CombineFileInputFormat$1$1.setInputFormat(CombineFileInputFormat.java:115)
    at org.apache.hadoop.mapreduce.JobSubmitter.writeSplits(JobSubmitter.java:306)
    at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:323)
    at org.apache.hadoop.mapreduce.Job$10.run(Job.java:200)
    at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1287)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:422)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1673)
    at org.apache.hadoop.mapreduce.Job.submit(Job.java:1287)
```

### Cause Analysis

According to the error stack, the memory overflows when the HDFS files are read during task submission. Generally, the memory is insufficient because the task needs to read a large number of small files.

### Solution

- Step 1** Check whether multiple HDFS files need to be read for the started MapReduce tasks. If yes, reduce the file quantity by merging the small-sized files in advance or using **combineInputFormat**.
- Step 2** Increase the memory when the **hadoop** command is run. The memory is set on the client. Change the value of **-Xmx** in **CLIENT\_GC\_OPTS** in the *Client installation directory/HDFS/component\_env* file to a larger value, for example, **512m**. Run the **source component\_env** command for the modification to take effect.

```
export YARN_ROOT_LOGGER=INFO,console
#GC_OPTS for client operation.
CLIENT_GC_OPTS="-Xmx512m Djava.io.tmpdir=${HADOOP_HOME}"
export HADOOP_CLIENT_OPTS="$CLIENT_GC_OPTS"
```

----End

## 23.3 Disk Space of a Node Is Used Up Due to Oversized Aggregated Logs of Yarn

### Issue

The disk usage of the cluster is high.

### Symptom

- On the host management page of Manager, the disk usage is too high.
- Only a few tasks are running on the Yarn web UI.

Cluster Metrics				
Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running
9	0	1	8	1
Cluster Nodes Metrics				
Active Nodes	Decommissioning Nodes	Decommissioned Nodes		
2	0	0		
Scheduler Metrics				
Scheduler Type	Scheduling Resource Type		Minimum Allocation	
Capacity Scheduler	(memory-mb (unit=M), vcores)		<memory:512, vCores:1>	
Show 20 entries				

- After the `hdfs dfs -du -h /` command is executed on the master node of the cluster, the command output shows that the following files consume a large amount of disk space.

```
22.5 G 45.0 G /tmp/logs/root/logs/application_1589278244866_0153
18.4 M 36.8 M /tmp/logs/root/logs/application_1589278244866_0154
23.4 G 46.8 G /tmp/logs/root/logs/application_1589278244866_0155
23.5 G 46.9 G /tmp/logs/root/logs/application_1589278244866_0156
23.7 G 47.4 G /tmp/logs/root/logs/application_1589278244866_0157
23.7 G 47.4 G /tmp/logs/root/logs/application_1589278244866_0158
22.5 G 45.0 G /tmp/logs/root/logs/application_1589278244866_0159
18.5 M 37.0 M /tmp/logs/root/logs/application_1589278244866_0160
22.5 G 45.0 G /tmp/logs/root/logs/application_1589278244866_0161
18.8 M 37.6 M /tmp/logs/root/logs/application_1589278244866_0162
24.0 G 48.0 G /tmp/logs/root/logs/application_1589278244866_0163
121.3 K 242.7 K /tmp/logs/root/logs/application_1589278244866_0164
1.1 M 2.1 M /tmp/logs/root/logs/application_1589278244866_0165
1.1 M 2.1 M /tmp/logs/root/logs/application_1589278244866_0166
1.1 M 2.1 M /tmp/logs/root/logs/application_1589278244866_0167
1.1 M 2.1 M /tmp/logs/root/logs/application_1589278244866_0168
```

- The log aggregation configuration parameters of the Mapreduce service is as follows.

* yarn.log-aggregation.retain-check-interval-seconds	86400
* yarn.log-aggregation.retain-seconds	1296000

## Cause Analysis

Jobs are submitted too frequently, and the time for deleting aggregated log files is set to 1296000, that is, aggregated logs are retained for 15 days. As a result, aggregated logs cannot be released within a short period of time, exhausting the disk space.

## Procedure

**Step 1** Log in to FusionInsight Manager and go to the Yarn configuration parameter page.

- MRS Manager: Log in to MRS Manager, choose **Services > Mapreduce > Service Configuration**, and select **All** from the **Type** drop-down list.
- FusionInsight Manager: Log in to FusionInsight Manager and choose **Cluster > Services > Mapreduce**. On the Mapreduce page, choose **Configurations > All Configurations**.



- Step 2** Search for the **yarn.log-aggregation.retain-seconds** parameter and decrease its value based on site requirements, for example, to **259200**. In this case, the aggregated logs are retained for three days, and the disk space is automatically released after the retention period expires.
- Step 3** Click **Save Configuration**. If a dialog box is displayed, deselect **Restart the affected services or instances**.
- Step 4** Restart the service whose configuration has expired during off-peak hours. The restart will interrupt upper-layer services and affect cluster management, maintenance, and services.
1. Log in to Manager.
  2. Restart the MapReduce and Yarn services.
- End

## 23.4 Temporary Files Are Not Deleted When a MapReduce Job Is Abnormal

### Issue

Temporary files are not deleted when a MapReduce job is abnormal.

### Symptom

There are too many files in the HDFS temporary directory, occupying too much memory.

### Cause Analysis

When a MapReduce job is submitted, related configuration files, JAR files, and files added by the **-files** parameter are uploaded to the temporary directory on HDFS so that the started container can obtain the files. The configuration item **yarn.app.mapreduce.am.staging-dir** determines the storage path. The default value is **/tmp/hadoop-yarn/staging**.

After a properly running MapReduce job is complete, temporary files are deleted. However, when a Yarn task corresponding to the job exits abnormally, temporary files are not deleted. As a result, the number of files in the temporary directory increases over time, occupying more and more storage space.

### Procedure

- Step 1** Log in to the cluster client.
1. Log in to any master node as user **root**. The user password is the one defined during cluster creation.
  2. If Kerberos authentication is enabled for the cluster, run the following commands to go to the client installation directory and configure environment variables. Then, authenticate the user and enter the password as prompted. Obtain the password from an administrator.  
**cd Client installation directory**

```
source bigdata_env
```

```
kinit hdfs
```

3. If Kerberos authentication is not enabled for the cluster, run the following commands to switch to user **omm** and go to the client installation directory to configure environment variables:

```
su - omm
```

```
cd Client installation directory
```

```
source bigdata_env
```

**Step 2** Obtain the file list.

```
hdfs dfs -ls /tmp/hadoop-yarn/staging/*/.staging/ | grep "^drwx" | awk '{print $8}' > job_file_list
```

The **job\_file\_list** file contains the folder list of all jobs. The following shows an example of the file content:

```
/tmp/hadoop-yarn/staging/omm/.staging/job_<Timestamp>_<ID>
```

**Step 3** Collect statistics on running jobs.

```
mapred job -list 2>/dev/null | grep job_ | awk '{print $1}' > run_job_list
```

The **run\_job\_list** file contains the IDs of running jobs. The content format is as follows:

```
job_<Timestamp>_<ID>
```

**Step 4** Delete running tasks from the **job\_file\_list** file to ensure that data of the running tasks is not deleted by mistake during deletion of expired data.

```
cat run_job_list | while read line; do sed -i "/$line/d" job_file_list; done
```

**Step 5** Delete expired data.

```
cat job_file_list | while read line; do hdfs dfs -rm -r $line; done
```

**Step 6** Delete temporary files.

```
rm -rf run_job_list job_file_list
```

```
----End
```

## 23.5 Incorrect Port Information of the Yarn Client Causes Error "connection refused" After a Task Is Submitted

### Issue

The ResourceManager of Yarn that requests to submit jobs throws error "connection refused", and the port number configured for Yarn is 8032.

## Symptom

One of Yarn's ResourceManager nodes in the MRS cluster cannot be connected, and the port number configured for Yarn is 8032.

## Cause Analysis

The service application runs outside the cluster, and the in-use client does not match the latest client configuration provided by the MRS cluster. The Yarn port is 8032, which is different from the actual port of Yarn's ResourceManager of MRS. As a result, the ResourceManager of Yarn that requests to submit jobs reports error "connection refused".

## Procedure

**Step 1** Update the MRS client.

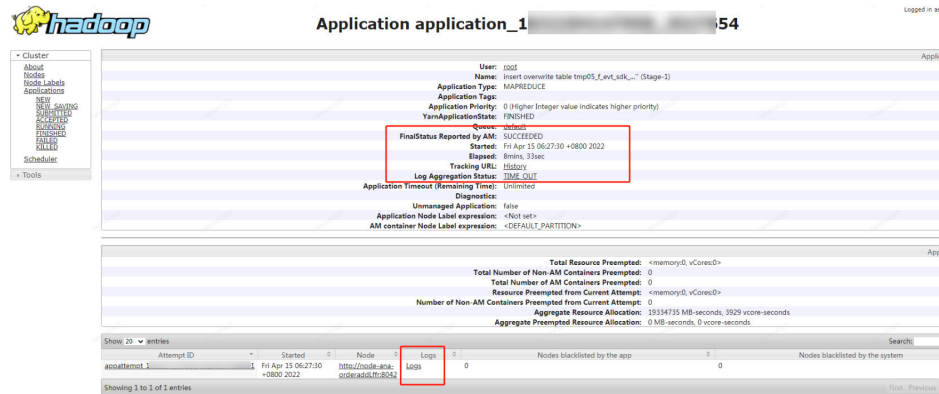
**Step 2** Submit the job again.

----End

# 23.6 "Could not access logs page!" Is Displayed When Job Logs Are Queried on the Yarn Web UI

## Symptom

When a user logs in to the Yarn web UI to view job logs and clicks **Local logs**, error message "Could not access logs page!" is displayed.





## Cause Analysis

**Local logs** is used to access service logs. However, for security purposes, this function is inaccessible from the Yarn web UI. The user can log in to the active ResourceManager node to view ResourceManager logs.

## Procedure

- Step 1** Log in to Manager and choose **Cluster > Services > Yarn**. On the **Yarn** page, click the **Instance** tab and take note of the service IP address of the active ResourceManager instance.
- Step 2** Log in to the active ResourceManager node as user **root**.
- Step 3** Go to the `/var/log/Bigdata/yarn/rm` directory and view the ResourceManager logs.

```
cd /var/log/Bigdata/yarn/rm
```

----End

## 23.7 Error "ERROR 500" Is Displayed When Queue Information Is Queried on the Yarn Web UI

### Symptom

When Yarn uses the Capacity scheduler, **ERROR 500** is reported after a user clicks a queue name on the Yarn web UI.

```
HTTP ERROR 500 javax.servlet.ServletException: javax.servlet.ServletException: java.lang.IllegalArgumentException: Illegal character in queryat index 81: https://XXX:20026/Yarn/ResourceManager/21/cluster/scheduler?openQueues= ^default$
```

### Cause Analysis

Symbol ^ in the URL cannot be identified. As a result, the page access fails.

## Procedure

- Step 1** Log in to Manager and choose **Cluster > Services > Yarn > Configurations > All Configurations**.
  - Step 2** Search for **yarn.resourcemanager.webapp.pagination.enable** in the search box.
  - Step 3** If the value is **true** (default), change it to **false** and save the configuration.
  - Step 4** On the Yarn page, click **Instance**, select all ResourceManager instances, click **More**, and select **Instance Rolling Restart**. Wait until the instances are started.
- End

## 23.8 Error "ERROR 500" Is Displayed When Job Logs Are Queried on the Yarn Web UI

### Symptom

A user views full logs on the Yarn web UI after a Spark Streaming job submitted on MRS has been running for a period of time. Error message "HTTP ERROR 500 org.apache.http.ConnectionCloseException: Premature end of chunk coded message body:closing chunk expected" is displayed.

### Cause Analysis

Because the job has been running for a long time, the size of logs to be displayed on the Yarn web UI is too large. In this case, you need to decrease the size of the aggregated log generated by the job so that the log can be displayed by segment.

## Procedure

- Step 1** Log in to the node where the Spark2x/Spark client is installed as user **root**.
  - Step 2** Run the following command to edit the file:  

```
vim $SPARK_HOME/conf/log4j-executor.properties
```
  - Step 3** Set **log4j.appender.sparklog.MaxFileSize** to a smaller value, for example, **20MB**. This parameter indicates the maximum size for storing a log file, and its default value is **50MB**.
  - Step 4** Set **log4j.appender.sparklog.MaxBackupIndex** to a smaller value, for example, **5**. This parameter indicates the maximum number of files that can be saved in rolling mode, and its upper limit is **10**. The eleventh file to be saved will overwrite the first one.
  - Step 5** Save the file.
  - Step 6** Submit the job again. The job runs properly.
- End

## 23.9 An Error Is Reported When a Yarn Client Command Is Used to Query Historical Jobs

### Symptom

When a Yarn client command is executed to query historical jobs, the following error is reported and the process is terminated.

```
[root@node-master1mup HDFS]# yarn app -list -appStates ALL
WARNING: YARN_CONF_DIR has been replaced by HADOOP_CONF_DIR. Using value of YARN_CONF_DIR.
WARNING: YARN_ROOT_LOGGER has been replaced by HADOOP_ROOT_LOGGER. Using value of YARN_ROOT_LOGGER.
#
# java.lang.OutOfMemoryError: Java heap space
# -XX:OnOutOfMemoryError="kill -9 %p"
#   Executing /bin/sh -c "kill -9 8642"...
sh: kill -9 8642: command not found
Error: A JNI error has occurred, please check your installation and try again
Exception in thread "main" java.lang.OutOfMemoryError: Java heap space
    at java.util.Arrays.copyOf(Arrays.java:3236)
    at sun.misc.Resource.getBytes(Resource.java:117)
    at java.net.URLClassLoader.defineClass(URLClassLoader.java:463)
    at java.net.URLClassLoader.access$100(URLClassLoader.java:74)
    at java.net.URLClassLoader$1.run(URLClassLoader.java:369)
    at java.net.URLClassLoader$1.run(URLClassLoader.java:363)
    at java.security.AccessController.doPrivileged(Native Method)
    at java.net.URLClassLoader.findClass(URLClassLoader.java:362)
    at java.lang.ClassLoader.loadClass(ClassLoader.java:418)
    at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:352)
    at java.lang.ClassLoader.loadClass(ClassLoader.java:351)
    at sun.launcher.LauncherHelper.checkAndLoadMain(LauncherHelper.java:601)
[root@node-master1mup HDFS]# hdfs dfs -ls /
#
# java.lang.OutOfMemoryError: GC overhead limit exceeded
# -XX:OnOutOfMemoryError="kill -9 %p"
#   Executing /bin/sh -c "kill -9 26409"...
sh: kill -9 26409: command not found
Exception in thread "main" java.lang.OutOfMemoryError: GC overhead limit exceeded
    at java.lang.String.substring(String.java:1969)
    at java.net.URLStreamHandler.parseURL(URLStreamHandler.java:254)
    at sun.net.www.protocol.file.Handler.parseURL(Handler.java:67)
    at java.net.URL.<init>(URL.java:640)
    at java.net.URL.<init>(URL.java:508)
    at sun.misc.URLClassPath$JarLoader.parseClassPath(URLClassPath.java:1232)
```

### Cause Analysis

The memory allocated to the client is insufficient. As a result, an error is reported when the command is submitted.

### Procedure

- Step 1** Log in to the node where the HDFS client is installed as user **root**.
- Step 2** Run the following command to edit the file:  
**vim /opt/client/HDFS/component\_env**
- Step 3** Increase the value of **export GC\_OPTS\_HDFS**, for example, to **Xmx512m**.
- Step 4** Save the file.
- Step 5** Run the following command to reload the environment variables:  
**source bigdata\_env**  
**----End**

## 23.10 Number of Files in the TimelineServer Directory Reaches the Upper Limit

### Symptom

In an MRS 3.x cluster, ResourceManager logs show that the number of TimelineServer data directories reaches the upper limit and a large number of error logs are printed.

The exception log is as follows:

```
The directory item limit of /tmp/hadoop-omm/yarn/timeline/generic-history/ApplicationHistoryDataRoot is exceeded: limit=1048576 items=1048576
```

### Cause Analysis

In MRS 3.x, TimelineServer uses an HDFS directory (for example, the **/tmp/hadoop-omm/yarn/timeline/generic-history/ApplicationHistoryDataRoot** directory in the preceding error information) to store historical task information. As a result, files in this directory accumulate until the number of directories reaches the upper limit configured in HDFS (the default value of **dfs.namenode.fs-limits.max-directory-items** is **1048576**).

In this case, set **yarn.timeline-service.generic-application-history.enabled** to **false** to obtain app task data from ResourceManager. This parameter specifies whether the client obtains app task data directly from TimelineServer.

### Procedure

- Step 1** Log in to FusionInsight Manager and choose **Cluster > Services > Yarn > Configurations > All Configurations**.
- Step 2** In the navigation pane on the left, choose **Yarn(Service) > Customization**. Locate the **yarn.yarn-site.customized.configs** parameter in the right pane, set the parameter name to **yarn.timeline-service.generic-application-history.enabled** and its value to **false**, and click **Save**.
- Step 3** Rolling restart the ResourceManager and TimelineServer instances.  
Click the **Instance** tab of the Yarn service, select all ResourceManager and TimelineServer instances, click **More**, and select **Instance Rolling Restart**.
- Step 4** (Optional) Rolling restart NodeManagers during off-peak hours based on service requirements.
- Step 5** After the instances are restarted, delete the directory for storing historical task information from HDFS, for example, **/tmp/hadoop-omm/yarn/timeline/generic-history/ApplicationHistoryDataRoot**.
  1. Log in to the client installation directory as the client installation user and configure environment variables.  

```
cd Client installation directory  
source bigdata_env
```

2. Run the following command to authenticate the user (skip this step for the user with Kerberos authentication disabled):

```
kinit Service user
```

3. Run the following command to delete the directory from HDFS:

```
hdfs dfs -rm -r /tmp/hadoop-omm/yarn/timeline/generic-history/  
ApplicationHistoryDataRoot/
```

**----End**



# 24 Using ZooKeeper

---

## 24.1 An Error Is Reported When the MRS Client Is Used to Connect to ZooKeeper

### Issue

An error is reported when a user attempts to access ZooKeeper from an MRS cluster.

### Symptom

When a user uses **zkcli.sh** to access ZooKeeper on the MRS master node, an error is reported.

### Cause Analysis

The command used by the user is incorrect. As a result, an error is reported.

### Procedure

**Step 1** Obtain the ZooKeeper IP address and port number. For details, see [How Do I Obtain the ZooKeeper Address?](#)

**Step 2** Log in to the master node as user **root**.

**Step 3** Run the following command to initialize environment variables:

```
source /opt/client/bigdata_env
```

**Step 4** Run the following command to connect to ZooKeeper:

```
zkCli.sh -server IP address of the ZooKeeper node:Port number
```

The IP address of the node where ZooKeeper is located is the one queried in [Step 1](#). Use commas (,) to separate multiple IP addresses.

**Step 5** Run common commands such as **ls /** to view ZooKeeper information.

----End

## 24.2 ZooKeeper Is Unavailable Because of Non-synchronized Time Between Active and Standby Master Nodes

### Symptom

The ZooKeeper service of the MRS cluster cannot be started, and an alarm is generated indicating that the ZooKeeper service is unavailable.

### Cause Analysis

Log in to the active and standby master nodes as user **root** and run the **ntpq -p** command. The command output shows that the time on the two nodes is not synchronized.

```
[root@node-master1u0ko ~]# ntpq -p
      remote           refid      st t when poll reach  delay  offset  jitter
=====
100.125.0.251  10.230.52.36   3 u   3  16  377   0.312  43.253  0.758
*LOCAL(0)      .LOCL.         0 l   -  16  377   0.000   0.000  0.000
[root@node-master1u0ko ~]# date
Tue Apr 27 09:12:36 CST 2021
```

### Procedure

- Step 1** Log in to the active OMS node as user **root** and run the **ifconfig** command to query the floating IP address of the current node.

In the following command output, for example, the floating IP address is 192.168.0.3.

```
[root@node-master1u0ko ~]# ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.176 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:11:d1:cf txqueuelen 1000 (Ethernet)
    RX packets 381241289 bytes 83002817678 (77.3 GiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 383655523 bytes 132910473015 (123.7 GiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

eth0:wsom: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.3 netmask 255.255.255.0 broadcast 192.168.0.255
    ether fa:16:3e:11:d1:cf txqueuelen 1000 (Ethernet)

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
    inet 127.0.0.1 netmask 255.0.0.0
    loop txqueuelen 0 (Local Loopback)
    RX packets 1722974128 bytes 660620647435 (615.2 GiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 1722974128 bytes 660620647435 (615.2 GiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
```

- Step 2** Log in to the master node where the NTP time is abnormal and run the following command to stop the NTP service:

```
service ntpd stop
```

- Step 3** Run the following command to manually synchronize the time between the active and standby master nodes:

```
/usr/sbin/ntpdate Floating IP address obtained in Step 1
```

**Step 4** Run the following command to start NTP:

```
service ntpd start
```

**Step 5** Run the **ntpstat** command to check the time synchronization result.

```
[root@node-master1u0ko ~]# ntpstat  
synchronised to local net at stratum 1  
time correct to within 7948 ms  
polling server every 16 s  
[root@node-master1u0ko ~]# date  
Tue Apr 27 09:14:53 CST 2021
```

----End

# 25 Storage-Compute Decoupling

---

## 25.1 A User Without the Permission on the /tmp Directory Failed to Execute a Job for Accessing OBS

### Symptom

When multiple MRS users access OBS and execute Spark, Hive, and Presto jobs, an error message is displayed, indicating that a user does not have the permission to access the `/tmp` directory.

### Cause Analysis

A temporary directory exists during job execution. The user who submits the job does not have the operation permission on the temporary directory.

### Procedure

**Step 1** On the **Dashboard** tab page of the cluster, query and record the name of the agency bound to the cluster.

**Step 2** Log in to the IAM console.

**Step 3** Choose **Permissions**. On the displayed page, click **Create Custom Policy**.

- **Policy Name:** Enter a policy name.
- **Scope:** Select **Global services**.
- **Policy View:** Select **Visual editor**.
- **Policy Content:**
  - a. **Allow:** Select **Allow**.
  - b. **Select service:** Select **Object Storage Service (OBS)**.
  - c. **Select action:** Select **WriteOnly**, **ReadOnly**, and **ListOnly**.
  - d. **Specific resources:**
    - i. Set **object** to **Specify resource path**, click **Add resource path**, and enter `obs_bucket_name/tmp/` and `obs_bucket_name/tmp/*` in **Path**.

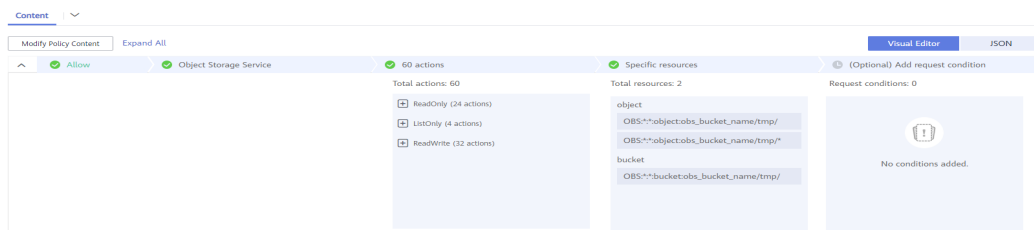
The `/tmp` directory is used as an example. If you need to add permissions for other directories, perform the following steps to add the directories and resource paths of all objects in the directories.

- ii. Set **bucket** to **Specify resource path**, click **Add resource path**, and enter `obs_bucket_name` in **Path**.

Replace `obs_bucket-name` with the actual OBS bucket name. If the bucket type is Parallel File System, you need to add the `obs_bucket_name/tmp/` path. If the bucket type is Object Storage, you do not need to add the path.

- e. (Optional) Request condition, which does not need to be added currently.

**Figure 25-1** Custom policy



**Step 4** Click **OK**.

**Step 5** Select **Agency** and click **Assign Permissions** in the **Operation** column of the agency queried in [Step 1](#).

**Step 6** Query and select the created policy in [Step 3](#).

**Step 7** Click **OK**.

----End

## 25.2 When the Hadoop Client Is Used to Delete Data from OBS, It Does Not Have the Permission for the .Trash Directory

### Symptom

After the `hadoop fs -rm obs://<obs_path>` command is executed, the following error information is displayed:

```
exception [java.nio.file.AccessDeniedException: user/root/.Trash/Current/: getFileStatus on user/root/.Trash/Current/: status [403]
```

### Cause Analysis

When deleting a file, Hadoop moves the file to the `.Trash` directory. If the user does not have the permission on the directory, error 403 is reported.

### Procedure

Solution 1:

Run the **hadoop fs -rm -skipTrash** command to delete the file.

Solution 2:

Add the permission to access the **.Trash** directory to the agency corresponding to the cluster.

**Step 1** On the **Dashboard** tab page of the cluster, query and record the name of the agency bound to the cluster.

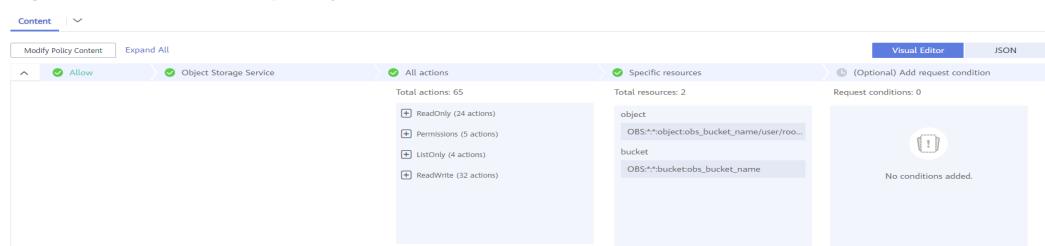
**Step 2** Log in to the IAM console.

**Step 3** Choose **Permissions**. On the displayed page, click **Create Custom Policy**.

- **Policy Name:** Enter a policy name.
- **Scope:** Select **Global services**.
- **Policy View:** Select **Visual editor**.
- **Policy Content:**
  - a. **Allow:** Select **Allow**.
  - b. **Select service:** Select **Object Storage Service (OBS)**.
  - c. Select all operation permissions.
  - d. **Specific resources:**
    - i. Set **object** to **Specify resource path**, click **Add resource path**, and enter the **.Trash** directory, for example, **obs\_bucket\_name/user/root/.Trash/\*** in **Path**.
    - ii. Set **bucket** to **Specify resource path**, click **Add resource path**, and enter **obs\_bucket\_name** in **Path**.
  - e. (Optional) Request condition, which does not need to be added currently.

Replace *obs\_bucket-name* with the actual OBS bucket name.

**Figure 25-2** Custom policy



**Step 4** Click **OK**.

**Step 5** Select **Agency** and click **Assign Permissions** in the **Operation** column of the agency queried in **Step 1**.

**Step 6** Query and select the created policy in **Step 3**.

**Step 7** Click **OK**.

**Step 8** Run the **hadoop fs -rm obs://<obs\_path>** command again.

----End

## 25.3 An MRS Cluster Fails Authentication When Accessing OBS Because the NTP Time of Cluster Nodes Is Not Synchronized

### Symptom

A 403 error is reported when the cluster accesses OBS.

### Cause Analysis

The NTP time of the master node in the cluster is not synchronized with the NTP server time of the nodes outside the cluster. The time difference exceeds 15 minutes. As a result, the authentication fails when the cluster accesses OBS and error 403 is reported.

### Procedure

**Step 1** Log in to the active Master node of the cluster as user **root**.

**Step 2** Modify the `${BIGDATA_HOME}/om-server/om/inst/conf/oms-config.ini` and `${BIGDATA_HOME}/om-server/OMS/workspace0/conf/oms-config.ini` configuration files, add `ntp.myhuaweicloud.com` before the value of `ntp_server_ip`.

```
...  
ntp_server_ip=ntp.myhuaweicloud.com,10.127.1.0  
...
```

**Step 3** Log in to the standby master node of the cluster as user **root** and go to [Step 2](#).

**Step 4** Wait about 2 minutes and run the following command on the active master node to check whether the IP address is successfully synchronized:

```
ntpq -np
```

The command output is as follows.

```
remote           refid           st t when poll reach  delay  offset  jitter  
-----  
* .127.1.0        .LOCL.         0 l  10  16  377  0.000  +0.000  0.000  
100.79. .         10.44. .       3 u  13  16  377  0.394  +47.857  0.280
```

----End